UNIVERSITY OF CALIFORNIA

SANTA CRUZ

**MEDIUM ACCESS CONTROL IN AD HOC NETWORKS WITH
OMNI-DIRECTIONAL AND DIRECTIONAL ANTENNAS**

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER ENGINEERING

by

Yu Wang

June 2004

The Dissertation of Yu Wang
is approved:

_____

Professor J.J. Garcia-Luna-Aceves, Chair

_____

Professor Anujan Varma

_____

Professor Roberto Manduchi

_____

Robert C. Miller
Vice Chancellor for Research and
Dean of Graduate Studies

Report Documentation Page

1. REPORT DATE **JUN 2004**	2. REPORT TYPE	3. DATES COVERED **00-06-2004 to 00-06-2004**

4. TITLE AND SUBTITLE **Medium Access Control in Ad Hoc Networks With Omni-Directional and Directional Antennas**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of California at Santa Cruz,Department of Computer Engineering,Santa Cruz,CA,95064**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **167**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Contents

List of Figures

List of Tables

<div align="center">**Abstract**</div>

<div align="center">Medium Access Control in Ad Hoc Networks with</div>

<div align="center">Omni-directional and Directional Antennas</div>

<div align="center">by</div>

<div align="center">Yu Wang</div>

An ad hoc network is a dynamic network formed on demand by a group of nodes without the aid of any pre-existing network infrastructure. An efficient and effective medium access control (MAC) protocol which regulates nodes' access to the shared channel(s) is essential in an ad hoc network. Our work is focused on the throughput and fairness properties of existing omni-directional MAC protocols as well as enhancement of their performance with directional antennas via both analytical and simulation approaches.

In the first part, we present the first analytical modeling of collision avoidance MAC protocols including the popular IEEE 802.11 MAC protocol in multi-hop ad hoc networks. We show that in ad hoc networks with a lot of hidden terminals, collision avoidance even if done correctly, can still limit achievable throughput significantly because of the much reduced spatial reuse. Then we advance the analytical modeling to evaluate those MAC protocols which use directional antennas and can achieve much higher throughput through directing transmissions and receptions to desired directions only. We show that the gain in spatial reuse outweighs that of collision avoidance and hence an aggressive all-directional scheme is more advantageous than other hybrid schemes that take conservative (or unnecessary) tradeoff between collision avoidance and spatial reuse. All the analytical work has provided very useful

insight on the interaction between spatial reuse, interference reduction and collision avoidance that previous work lacked.

The work done in the first part also reveals the fairness problem inherent in IEEE 802.11 based wireless networks which can hinder the deployment of high profile applications that require some quality of service (QoS) assurances. In the second part of the thesis work, we first propose a hybrid channel access scheme that can achieve better fairness while maintaining compatibility with the existing IEEE 802.11 standard. Then we propose a fairness framework to address the problem as well as mechanisms to realize the framework and show that the resulting scheme can achieve far better fairness than previous schemes with only moderate throughput degradation.

Acknowledgements

First of all, I want to express my deep gratitude to my advisor, Professor J. J. Garcia-Luna-Aceves for his continuous support, guidance, encouragement and understanding during my journey of doctoral study and research. Besides JJ's professional mentorship, I am also especially indebted for JJ's support during my stay in Paris, France in the year 2002 and three vacations back to China in the year 2003 for family reunion, marriage and visa which will be always remembered.

Second, I thank the members of my committee, Professors Anujan Varma and Roberto Manduchi for their valuable advice and feedback. I also thank Professors Glen Langdon and Charlie McDowell for sitting on my qualifying exam committee and Professors Brahim Bensaou (Hong Kong University of Science and Technology), K. C. Chua (National University of Singapore), Katia Obraczka, Jim Whitehead and Xiubin Zhang (Shanghai Jiao-tong University) for their generous help when I need it.

Third, I want to thank my fellow *cocos* (Chane, Soumya, Lori, Lichun, Marco, Marcelo Carvalho, Marc, Hong, Long, Yong, Zhenjiang, Lei, Ravi, Saro, Hari, Ramesh, Brad, Renato and Radhika) in the Computer Communication Research Group (CCRG) directed by JJ and students (Nacho, Cintia, Kumar and Venkatesh) in Katia's lab for their fellowship. I also thank the staff in School of Engineering (Carol Mullane, Jodi Rieger, Frances Ruiz-Wood, Haifang Telc, Ma Xiong and more) for their support. I have indeed enjoyed working in such a congenial environment in the past four years.

Fourth, I want to thank my friends and their families both in and out of school whom I have come to know throughout the years, including Miangang Zhao, Wei Gong, Zheng Xu,

To my parents and Jin

Chapter 1

Introduction

An ad hoc network is a dynamic network formed on demand by a group of nodes without the aid of any pre-existing network infrastructure. It can be deployed rapidly in disaster relief, battlefield communications, collaborative computing and other applications and has received increasing interest in recent years [1–10]. Due to the scarce wireless resource available, an efficient and effective medium access control (MAC) protocol which regulates nodes' access to the shared channel(s) is essential in an ad hoc network. In the recent past, a lot of MAC protocols have been proposed and they can be largely divided into two class: Contention-based and schedule-based. In contention-based MAC protocols (e.g., [11–24]), nodes' access to the shared channel are not synchronized, and they contend for the channel whenever there are packets in their buffers ready to be sent. This class of MAC protocols is attractive due to its simpleness, robustness and suitability for bursty-traffic. In the second class, nodes' access to the shared channel are synchronized and packets are transmitted in time-slots. Schedule-based MAC protocols (e.g., [25,26]) can achieve contention-free transmission after the necessary in-

formation for scheduling is exchanged among nodes and schedules are made. These protocols can usually achieve a much higher and stable throughput when the load to the shared channel is high despite the requirement for time-slotted structures and overhead in exchanging neighbor information. In this thesis work, we focus on the throughput and fairness properties of the existing contention-based MAC protocols as well as enhancement of their performance with directional antennas via both analytical and simulation approaches.

In contention-based MAC protocols collision avoidance is very important, because simple MAC protocols such as carrier sense multiple access (CSMA) cannot combat the "hidden terminal" problem and performance can degrade to that of the ALOHA protocol in ad hoc networks [11, 12, 27].

Many collision-avoidance protocols [13,14,16,17] have been proposed and the most popular collision avoidance scheme today consists of a four-way sender-initiated handshake in which the transmission of a data packet and its acknowledgment is preceded by request-to-send (RTS) and clear-to-send (CTS) packets between a pair of sending and receiving nodes.[1] Other nodes that overhear RTS or CTS packets will defer their access to the channel to avoid collisions. For the sake of simplicity, it can be also called RTS/CTS based scheme as illustrated in Figure 1.1. Among all these proposed collision-avoidance protocols, the IEEE 802.11 distributed foundation wireless medium access control (DFWMAC) protocol [16] is very popular in the performance studies of routing protocols for ad hoc networks, even though it was originally intended for wireless LANs with no or very few hidden terminals. Though there has been considerable work on the performance evaluation of IEEE 802.11 and simi-

---

[1] This is in contrast to receiver-initiated MAC schemes (e.g. [23, 24]) in which the collision avoidance handshake is started by receivers.

2
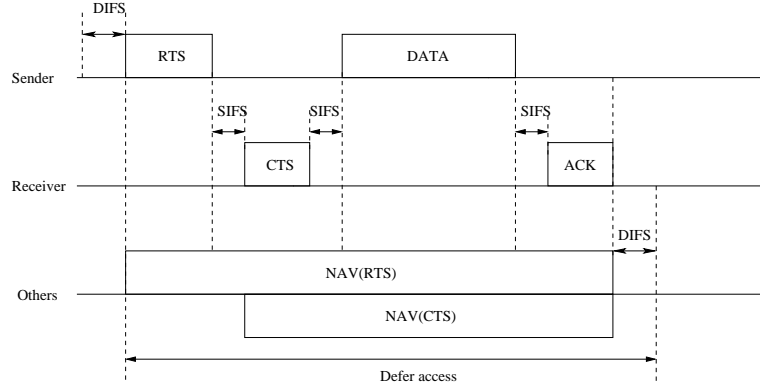
Figure 1.1: RTS/CTS based four-way collision avoidance handshake

lar protocols [28–34], most of the analytical models are largely confined to single-hop networks [32–34] or cases when the number of hidden terminals is very small [28].

We deem it very important to investigate the performance of the four-way sender-initiated collision avoidance scheme with a truly multi-hop network model as potential interference from hidden nodes always exists, which is a salient characteristic of multi-hop ad hoc networks. Hence, in Chapter 2, we present the first analytical model to derive the saturation throughput of four-way sender-initiated collision avoidance protocols in multi-hop ad hoc networks. We show that the sender-initiated collision-avoidance scheme achieves much higher throughput than the idealized carrier sense multiple access scheme with an ideal separate channel for acknowledgments. More importantly, we show that the collision avoidance scheme can accommodate much fewer competing nodes within a region in a network infested with hidden terminals than in a fully-connected network, if reasonable throughput is to be maintained. This shows that the scalability problem of contention-based collision-avoidance protocols looms much earlier than people might expect. Simulations of the popular IEEE 802.11 MAC protocol validate the predictions made in the analysis. Simulation results

3

also reveal the fairness problem in ad hoc networks which refers to the severe degradation in throughput experienced by some nodes due to location dependent contention and motivate our work on fairness in Chapters 5, 6 and 7.

In recent years, there has been also increasing interest in using directional antennas to improve the performance of existing MAC protocols [35–41]. Some MAC protocols using directional antennas have been proposed in the past, which achieve tradeoff between spatial reuse and collision avoidance via a combination of omni-directional and directional transmission modes. Simulation-based studies of these proposed protocols show that they have improved performance over the existing omni-directional IEEE 802.11 MAC protocol. However, the majority of the performance analyses of directional collision avoidance schemes have been done via simulations, and there is little prior work on the analytical modeling of directional collision avoidance protocols. Hence, in Chapter 3, we investigate the interaction between spatial reuse, interference reduction and collision avoidance by extending the model used in Chapter 2 to analyze several typical MAC protocols using directional antennas. In our model, we consider both (a) the effect of directional transmitting and receiving on spatial reuse and collision avoidance, and (b) the effect of the differences in gains between omni-directional and directional transmissions. Analytical results show that, when the directional collision avoidance scheme in which all transmissions are directional is augmented with directional receiving, one-hop throughput does not decrease due to the increased spatial reuse, even when the number of competing nodes within a region increases. This is very desirable because the scalability problem shown in Chapter 2 is mitigated by the use of powerful directional antenna systems. It is also shown that, as expected, the performance of directional

4

collision avoidance schemes degrades when directional transmissions have much higher gain than omni-directional transmissions. However, this degradation is relatively small. Simulations of the IEEE 802.11 protocol and its directional variants validate the results predicted in the analysis; and show that side lobes affect little on throughput if the gain of the main transmission lobe is reasonably higher than that of side lobes and the carrier sensing threshold is raised to make nodes less sensitive to channel activities.

In Chapter 4, we study via simulations both the performance of unicast traffic in the presence of broadcast traffic and the performance of broadcast traffic when mixed with unicast traffic, which is different from previous investigations reported in the literature in which broadcast traffic is investigated in isolation. The simulation results show that the presence of broadcast traffic does not degrade the performance of the all-directional collision avoidance scheme significantly, even for relatively large percentages of broadcast traffic. The work done in Chapters 3 and 4 indicates that the most attractive collision avoidance approach consists of using directional transmissions of control and data packets, together with the directional reception of packets whenever a node is expecting a particular packet. Given the high tolerance to broadcast traffic of directional collision avoidance schemes, it is argued that the periodic transmission of beacons omni-directionally suffices to provide such schemes with the relative locations of neighboring nodes.

Mitigating the fairness problems in multi-hop ad hoc networks has also been investigated intensively in the recent past [14, 42–50]. Fairness problems usually result from location-dependent contention, which is very common in multi-hop ad hoc networks. Additionally, the commonly used binary exponential backoff (BEB) scheme can aggravate the

5

fairness problem despite its robustness against repetitive collisions, because the node that suc-

ceeds in the last transmission period will gain access to the shared channel again with much

higher probability while other nodes are denied access almost completely. We investigate the

fairness problem in detail in Chapter 5. We show that the required multi-hop coordination

makes those backoff-based distributed fair queueing schemes less effective. Using extensive

simulations of two competing flows with different underlying network configurations, it is

shown that the commonly used flow contention graph is insufficient to model the contention

among nodes and that various degrees of unfairness can take place. The fairness problem

is more severe in TCP-based flows due to the required acknowledgment traffic, and TCP

throughput is also negatively affected.

In Chapter 6, we propose a novel hybrid channel access scheme that combines both

sender-initiated and receiver-initiated collision avoidance handshake. The new scheme is com-

patible with the popular IEEE 802.11 MAC protocol and involves only some additional queue

management and book-keeping work. Simulation experiments show that the new scheme can

alleviate the fairness problems existent for both UDP and TCP based applications with almost

no degradation in throughput. However, it still cannot solve the fairness problem conclusively.

This indicates that more explicit information exchange among contending nodes is mandatory

to solve the fairness problem which motivates the further work done in Chapter 7.

In Chapter 7, we introduce a framework to address the fairness problem in ad hoc

networks. The framework includes four key components: Exchanging flow contention in-

formation, using an adaptive backoff algorithm that does not aggravate the fairness problem

like the binary exponential backoff (BEB) does, introducing a combination of sender-initiated

and receiver-initiated collision avoidance handshake, and dealing with two-way flows. We explain the rationale for these components and then propose some algorithms to realize the framework. The resulting scheme, which we call topology aware fair access (TAFA) is compared through simulations with the IEEE 802.11 MAC protocol and the hybrid channel access scheme proposed in Chapter 6. Simulation results show that TAFA can solve the fairness problem in UDP-based applications with negligible degradation in throughput. It can also solve the notorious problem of the starvation of flows in TCP-based applications, while incurring only some throughput degradation. Hence, TAFA shows a much better overall tradeoff between throughput and fairness than other schemes previously proposed.

Papers [51–63] based on the above research work have been published or accepted for publication. Chapter 8 summarizes our contribution and proposes some future work.

Chapter 2

Omni-directional Medium Access

Control (MAC) Schemes

2.1 Introduction

As discussed in Chapter 1, many collision-avoidance protocols [13, 14, 16, 22] have

been proposed in the recent past to combat the hidden terminal problem for wireless networks.

Among all these protocols, the IEEE 802.11 MAC protocol [16] is the most popular which is

based on a RTS-CTS-data-ACK four-way collision avoidance handshake with non-persistent

carrier sensing. Since the inception of the IEEE 802.11 working group, there has been con-

siderable work on the performance evaluation of IEEE 802.11 in wireless LANs and possible

ways to enhance its performance [28–34, 64–69]. Some of them are simulation based, others

also use analytical models. However, the analytical models of collision-avoidance protocols

so far are largely confined to single-hop networks [32–34, 68] or cases when the number of

8

hidden terminals is very small [28, 65]. They cannot capture a salient feature of ad hoc networks, i.e., potential interference from hidden nodes always exists. Hence, we are interested in investigating the performance of the sender-initiated collision avoidance scheme based on a four-way handshake in a general framework that is more applicable to ad hoc networks. Our work is different from the work done by Gupta and Kumar [70] in that the authors give very general results about the capacity of wireless networks from an information-theory centric view, while our analysis and simulation experiments show how the typical collision-avoidance MAC protocols can perform in multi-hop ad hoc networks.

In Section 2.2, we present the analysis of the sender-initiated collision-avoidance scheme based on a four-way handshake and non-persistent carrier sensing, which can be also called the RTS/CTS-based scheme for the sake of simplicity. To our knowledge, this is the first analytical modeling of collision avoidance in multi-hop networks. We first adopt a simple model in which nodes are randomly placed on a plane according to two-dimensional Poisson distribution with density λ. Varying λ has the effect of changing the congestion level within a region as well as the number of hidden terminals. In this model, it is also assumed that each node is ready to transmit independently in each time slot with probability p, where p is a protocol-dependent parameter. This model was first used by Takagi and Kleinrock [71] to derive the optimum transmission range of a node in a multi-hop wireless network, and was used subsequently by Wu and Varshney [72] to derive the throughputs of non-persistent CSMA and some variants of busy tone multiple access (BTMA) protocols [27]. Then we assume that both carrier sensing and collision avoidance work perfectly, that is, that nodes can accurately sense the channel busy or idle, and that the RTS/CTS scheme can avoid the transmission of data

packets that collide with other packets at the receivers. The latter assumption can be called *perfect collision avoidance* and has been shown to be doable in the floor acquisition multiple access (FAMA) protocol [15, 17, 22]. Later we extend this model to take into account the possibility of data packets colliding with other transmissions, so that the model is also applicable to other MAC protocols, such as the popular IEEE 802.11 protocol, in which perfect collision avoidance is not strictly enforced.

In Section 2.3, we present numerical results from our analysis. We compare the performance of the sender-initiated collision avoidance scheme against the idealized non-persistent CSMA protocol in which a secondary channel is assumed to send acknowledgments in zero time and without collisions [27, 72], as the latter is the only protocol whose analysis for multi-hop ad hoc networks is available for comparison to date. It is shown that the RTS/CTS scheme can achieve far better throughput than the CSMA protocol, even when the overhead due to RTS/CTS exchange is high. The results illustrate the importance of enforcing collision avoidance in the RTS/CTS handshake. However, the analytical results also indicate that the aggregate throughput of sender-initiated collision avoidance drops faster than that in a fully-connected network when the number of competing nodes within a region increases. This contrasts with conclusions drawn from the analysis of collision avoidance in fully-connected networks or networks with limited hidden terminals [22]. Our results show that hidden terminals degrade the performance of collision avoidance protocols beyond the basic effect of having a longer vulnerability period for RTSs. Hence, it follows that collision avoidance becomes more and more ineffective for a relatively crowded region with hidden terminals.

To validate the findings drawn from this analysis, in Section 2.4 we present simulations of the popular IEEE 802.11 MAC protocol. The simulation results clearly show that the IEEE 802.11 MAC protocol cannot ensure collision-free transmission of data packets, and that almost half of the data packets transmitted cannot be acknowledged due to collisions, even when the number of competing nodes in a neighborhood is only eight! We also investigate a variant of the IEEE 802.11 MAC protocol in which the contention window used in deciding backoff time is fixed. This variant does not have the inherent fairness problem in the original binary exponential backoff (BEB) scheme used in the IEEE 802.11 MAC protocol, though it is not fine tuned to achieve the best performance. However, the simulation results do show that decreasing the contention window leads to more collisions of data packets, while increasing the contention window leads to more wasted time in waiting. Both approaches can limit the maximum achievable throughput significantly, which is a typical dilemma for contention-based MAC protocols, especially for those that do not provide *correct* collision avoidance scheme in less crowded multi-hop networks. The performance of the simulated IEEE 802.11 MAC protocol correlates well with what is predicted in the extended analysis, which takes into account the effect of data packet collisions and is used for the case when the number of competing nodes in a region is small. When the number of competing nodes in a region increases, the performance gap between IEEE 802.11 and the analysis decreases, which validates the statement that even a perfect collision-avoidance protocol loses its effectiveness gradually due to the random nature of the channel access and the limited information available to competing nodes.

Section 2.5 concludes this chapter with possible ways to improve the performance

of collision avoidance protocols in ad hoc networks and highlights the usage of directional antennas, which is analyzed subsequently in Chapters 3 and 4.

2.2   Approximate Analysis

In this section, we derive the approximate throughput of a perfect collision avoidance protocol. In our network model, nodes are two-dimensionally Poisson distributed over a plane with density λ, i.e., the probability $p(i, S)$ of finding i nodes in an area of S is given by:

$$p(i, S) = \frac{(\lambda S)^i}{i!} e^{-\lambda S}.$$

Assume that each node has the same transmission and receiving range of R, and denote by N the average number of nodes within a circular region of radius R; therefore, we have $N = \lambda \pi R^2$.

To simplify our analysis, we assume that nodes operate in time-slotted mode. As prior results for CSMA and collision-avoidance protocols show [27], the performance of MAC protocols based on carrier sensing is much the same as the performance of their time-slotted counterparts in which the length of a time slot equals one propagation delay and the propagation delay is much smaller than the transmission time of data packets.

The length of each time slot is denoted by τ. Note that τ is not just the propagation delay, because it also includes the overhead due to the transmit-to-receive turn-around time, carrier sensing delay and processing time. In effect, τ represents the time required for all the nodes within the transmission range of a node to know the event that occurred τ seconds ago. The transmission times of RTS, CTS, data, and ACK packets are normalized with regard to

12

τ, and are denoted by l_{rts}, l_{cts}, l_{data}, and l_{ack}, respectively. Thus, τ is also equivalent to 1 in later derivations. For the sake of simplicity, we also assume that all packet transmission times are multiples of the length of a time-slot.

We derive the protocol's throughput based on the heavy-traffic assumption, i.e., a node always has a packet in its buffer to be sent and the destination is chosen randomly from one of its neighbors. This is a fair assumption in ad hoc networks in which nodes are sending data and signaling packets continually. We also assume that a node is ready to transmit with probability p and not ready with probability $1 - p$. Here p is a protocol-specific parameter that is slot independent. At the level of individual nodes, the probability of being ready to transmit may vary from time slot to slot, depending on the current states of both the channel and the node. However, because we are interested in deriving the average performance metrics instead of instantaneous or short-term metrics, the assumption of a fixed probability p may be considered as an averaged quantity that can still reasonably approximate the factual burstiness from a long-term point of view. In fact, this assumption is necessary to make the theoretical modeling tractable and has been extensively applied before [34, 71, 72]. For example, this model was used by Takagi and Kleinrock [71] to derive the optimal transmission range of a node in a multi-hop wireless network, and was used subsequently by Wu and Varshney [72] to derive the throughput of non-persistent CSMA and some variants of busy tone multiple access (BTMA) protocols [27].

It should also be noted that, even when a node is ready to transmit, it may transmit or not in the slot, depending on the collision avoidance and resolution schemes being used, as well as the channel's current state. Thus, we are more interested in the probability that a node

13

transmits in a time slot, which is denoted by p'. Similar to the reasoning presented for p, we also assume that p' is independent at any time slot to make the analysis tractable. Given this simplification, p' can be defined to be

$$p' = p \cdot \text{Prob.\{Channel is sensed idle in a slot\}}$$

$$\approx p \cdot \Pi_I$$

where Π_I is the limiting probability that the channel is in idle state, which we derive subsequently.

We are not interested in the exact relationship between p and p', and it is enough to obtain the range of values that p' can take, because the throughput of these protocols is mostly influenced by p'. To derive the rough relationship between p and p', we set up a channel model that includes two key simplifying assumptions.

First, we model the channel as a circular region in which there are some nodes. The nodes within the region can communicate with each other while they have weak interactions with nodes outside the region. *Weak interaction* means that the decision of inner nodes to transmit, defer and back off is almost not affected by that of outer nodes and vice versa. Considering that nodes do not exchange status information explicitly (e.g., either defer due to collision avoidance or back off due to collision resolution), this assumption is reasonable and helps to simplify the model considerably. Thus, the channel's status is only decided by the successful and failed transmissions within the region.

Second, we still consider the failed handshakes initiated by nodes within the region to outside nodes, because this has a direct effect on the channel's usability for other nodes within the region. Though the radius of the circular region R' is unknown, it falls between

14

$R/2$ and $2R$. This follows from noting that the maximal radius of a circular region in which all nodes are guaranteed to hear one another equals $R' = R/2$, and all the direct neighbors and hidden nodes are included into the region when $R' = 2R$. Thus, we obtain $R' = \alpha R$ where $0.5 \leq \alpha \leq 2$, and α needs to be estimated.

With the above assumptions, the channel can be modeled by a four-state Markov chain illustrated in Figure 2.1. The significance of the states of this Markov chain is the following:

- *Idle* is the state when the channel around node x is sensed idle, and obviously its duration is τ.

- *Long* is the state when a successful four-way handshake is done. For simplicity, we assume that the channel is in effect busy for the duration of the whole handshake, thus the busy time T_{long} is

$$T_{long} = l_{rts} + \tau + l_{cts} + \tau + l_{data} + \tau + l_{ack} + \tau$$

$$= l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau.$$

- *Short1* is the state when multiple nodes around the channel transmit RTS packets during the same time slot and their transmissions collide. The busy time of the channel T_{short1} is therefore

$$T_{short1} = l_{rts} + \tau.$$

- *Short2* is the state when one node around the channel initiates a failed handshake with a node outside the region. Even though a CTS packet may not be sent due to the collision

15
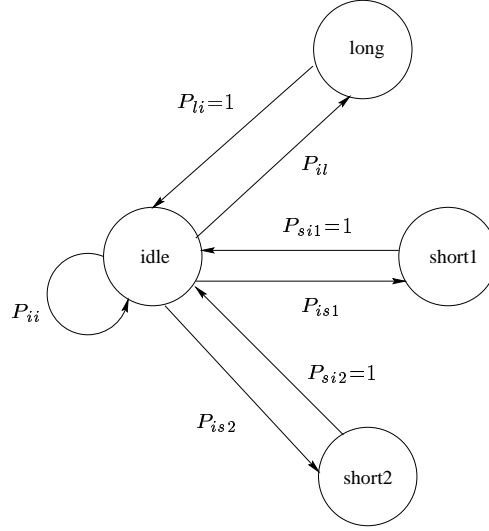
Figure 2.1: Markov chain model for the channel around a node

of the sending node's RTS packet with other packets originated from nodes outside the region or due to the deferring of the receiving node to other nodes, those nodes overhearing the RTS as well as the sending node do not know if the handshake is successfully continued, until the time required for receiving a CTS packet elapses. Therefore the channel is in effect busy, i.e., unusable for all the nodes sharing the channel, for the time stated below:

$$T_{short2} = l_{rts} + \tau + l_{cts} + \tau$$

$$= l_{rts} + l_{cts} + 2\tau.$$

Now we proceed to calculate the transition probabilities of the Markov chain.

In most collision avoidance schemes with non-persistent carrier sensing, no node is allowed to transmit immediately after the channel becomes idle, thus the transition probabilities from *long* to *idle*, from *short1* to *idle* and from *short2* to *idle* are all 1.

According to the Poisson distribution of the nodes, the probability of having i nodes

within the receiving range R of x is $e^{-N}N^i/i!$, where $N = \lambda\pi R^2$. Therefore, the mean number of nodes that belong to the shared channel is $M = \lambda\pi R'^2 = \alpha^2 N$. Assuming that each node transmits independently, the probability that none of them transmits is $(1 - p')^i$, where $(1 - p')$ is the probability that a node does not transmit in a time slot. Because the transition probability P_{ii} from *idle* to *idle* is the probability that none of the neighboring nodes of x transmits in this slot, P_{ii} is given by

$$
\begin{aligned}
P_{ii} &= \sum_{i=0}^{\infty} (1 - p')^i \frac{M^i}{i!} e^{-M} \\
&= \sum_{i=0}^{\infty} \frac{[(1 - p')M]^i}{i!} e^{-(1-p')M} \cdot e^{-p'M} \\
&= e^{-p'M}.
\end{aligned}
$$

We average the probabilities over the number of interfering nodes in a region because of two reasons. First, it is much more tractable than the approach that conditions on the number of nodes, calculates the desired quantities, and then uses the Poisson distribution to obtain the average. Second, in our simulation experiments, we fix the number of competing nodes in a region (which is N) and then vary the location of the nodes to approximate the Poisson distribution, which is configurationally closer to our analytical model; the alternative would be to generate $2, 3, 4, \ldots$ nodes within one region, get the throughput for the individual configuration and then calculate the average, which is not practical.

Next we need to calculate the transition probability P_{il} from *idle* to *long*. If there are i nodes around node x, for such a transition to happen, one and only one node should be able to complete one successful four-way handshake while other nodes do not transmit. Let p_s denote the probability that a node begins a successful four-way handshake at each slot, we

can then calculate P_{il} as follows:

$$P_{il} = \sum_{i=1}^{\infty} i p_s (1 - p')^{i-1} \frac{M^i}{i!} e^{-M}$$

$$= \sum_{i=1}^{\infty} p_s (1 - p')^{i-1} \frac{M^{i-1}}{(i-1)!} M e^{-M}$$

$$= p_s M \sum_{i=0}^{\infty} \frac{[M(1 - p')]^i}{i!} e^{-M(1-p'+p')}$$

$$= p_s M e^{-p'M}.$$

To obtain the above result, we use the fact that the distribution of the number of nodes within R' does not depend on the existence of node x, because of the memoryless property of the Poisson distribution. Up to this point, p_s is still an unknown quantity that we derive subsequently.

The transition probability from *idle* to *short1* is the probability that more than one node transmit RTS packets in the same slot; therefore, P_{is1} can be calculated as follows:

$$P_{is1} = \sum_{i=2}^{\infty} [1 - (1 - p')^i - i p' (1 - p')^{i-1}] \frac{M^i}{i!} e^{-M}$$

$$= 1 - (1 + M p') e^{-p'M}.$$

Having calculated P_{ii}, P_{il} and P_{is1}, we can calculate P_{is2}, the transition probability from *idle* to *short2*

$$P_{is2} = 1 - P_{ii} - P_{il} - P_{is1}$$

$$= 1 - e^{-p'M} - p_s M e^{-p'M} - (1 - (1 + M p') e^{-p'M})$$

$$= (p' - p_s) M e^{-p'M}.$$

Let π_i, π_l, π_{s1} and π_{s2} denote the steady-state probabilities of states *idle*, *long*, *short1* and

short2, respectively. From Figure 2.1, we have

$$\pi_i P_{ii} + \pi_l + \pi_{s1} + \pi_{s2} = \pi_i$$

$$\pi_i P_{ii} + 1 - \pi_i = \pi_i$$

$$\pi_i = \frac{1}{2 - P_{ii}} = \frac{1}{2 - e^{-p'M}}.$$

The limiting probability Π_I, i.e., the long run probability that the channel around node x is found idle, can be obtained by:

$$\Pi_I = \frac{\pi_i T_{idle}}{\pi_i T_{idle} + \pi_l T_{long} + \pi_{s1} T_{short1} + \pi_{s2} T_{short2}}.$$

Noting that $\pi_i P_{il} = \pi_l, \pi_i P_{is1} = \pi_{s1}$ and $\pi_i P_{is2} = \pi_{s2}$, we obtain

$$\Pi_I = \frac{\pi_i T_{idle}}{\pi_i T_{idle} + \pi_i P_{il} T_{long} + \pi_i P_{is1} T_{short1} + \pi_i P_{is2} T_{short2}}$$

$$= \frac{T_{idle}}{T_{idle} + P_{il} T_{long} + P_{is1} T_{short1} + P_{is2} T_{short2}}.$$

The relationship between p' and p is then:

$$p' = \frac{p T_{idle}}{T_{idle} + P_{il} T_{long} + P_{is1} T_{short1} + P_{is2} T_{short2}}$$

$$= \frac{p T_{idle}}{T_{idle} + p_s M e^{-p'M} T_{long} + (1 - (1 + p'M)e^{-p'M}) T_{short1} + (p' - p_s) M e^{-p'M} T_{short2}}$$

$$= \frac{p\tau}{\tau + p_s M e^{-p'M}(l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau) + (1 - (1 + p'M)e^{-p'M})(l_{rts} + \tau) + \cdots}$$

$$\overline{\cdots + (p' - p_s) M e^{-p'M}(l_{rts} + l_{cts} + 2\tau)}$$

(2.1)

In the above equation, the probability that a node x starts successfully a four-way handshake in a time slot, p_s, is yet to be determined.

The states of a node x can be modeled by a three-state Markov chain, which is shown in Figure 2.2. In Figure 2.2, *wait* is the state when the node defers for other nodes or

19
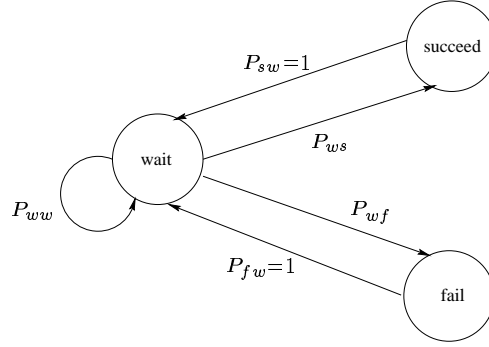
Figure 2.2: Markov chain model for a node

backs off, *succeed* is the state when the node can complete a successful four-way handshake with other nodes, and *fail* is the state when the node initiates an unsuccessful handshake. For simplicity, we regard *succeed* and *fail* as the states when two different kinds of *virtual* packets are transmitted and their lengths are:

$$T_{succeed} = T_{long}$$

$$= l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau$$

$$T_{fail} = T_{short2}$$

$$= l_{rts} + l_{cts} + 2\tau.$$

Obviously, the duration of a node in *wait* state T_{wait} is τ.

Because by assumption collision avoidance is enforced at each node, no node is allowed to transmit data packets continuously; therefore, the transition probabilities from *succeed* to *wait* and from *fail* to *wait* are both one.

To derive the transition probability P_{ws} from *wait* to *succeed*, we need to calculate the probability $P_{ws}(r)$ that node x successfully initiates a four-way handshake with node y at a given time slot when they are at a distance r apart. Before calculating $P_{ws}(r)$, we define $B(r)$ to be the area that is in the hearing region of node y but outside the hearing region of node x,
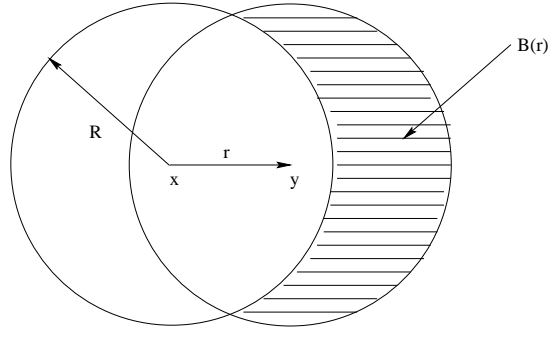
Figure 2.3: Illustration of the "hidden" area

i.e., the interfering region "hidden" from node x as the shaded area shown in Figure 2.3. $B(r)$ has been shown in [71] to be:

$$B(r) = \pi R^2 - 2R^2 q\left(\frac{r}{2R}\right) \tag{2.2}$$

where $q(t) = \arccos(t) - t\sqrt{1 - t^2}$.

Then $P_{ws}(r)$ can be calculated as:

$$P_{ws}(r) = P_1 \cdot P_2 \cdot P_3 \cdot P_4(r)$$

where

$P_1 = \text{Prob.}\{x \text{ transmits in a slot}\},$

$P_2 = \text{Prob.}\{y \text{ does not transmit in the time slot}\},$

$P_3 = \text{Prob.}\{\text{none of the terminals within } R \text{ of } x \text{ transmits in the same slot}\},$

$P_4(r) = \text{Prob.}\{\text{none of the terminals in } B(r) \text{ transmits for } (2l_{rts} + 1) \text{ slots} \mid r\}.$

The reason for the last term is that the vulnerable period for an RTS is only $2l_{rts} + 1$, and once the RTS is received successfully by the receiving node (which can then start sending the CTS), the probability of further collisions is assumed to be negligibly small.

21

Obviously, $P_1 = p'$ and $P_2 = (1 - p')$. On the other hand, P_3 can be obtained by

$$P_3 = \sum_{i=0}^{\infty} (1 - p')^i \frac{(\lambda \pi R^2)^i}{i!} e^{-\lambda \pi R^2}$$

$$= \sum_{i=0}^{\infty} (1 - p')^i \frac{N^i}{i!} e^{-N}$$

$$= e^{-p'N}.$$

Similarly, the probability that none of the terminals in $B(r)$ transmits in a time slot is given

by

$$p_4(r) = \sum_{i=0}^{\infty} (1 - p')^i \frac{(\lambda B(r))^i}{i!} e^{-\lambda B(r)}$$

$$= e^{-p' \lambda B(r)}.$$

Hence, $P_4(r)$ can be expressed as

$$P_4(r) = (p_4(r))^{2l_{rts}+1}$$

$$= e^{-p' \lambda B(r)(2l_{rts}+1)}.$$

Given that each sending node chooses any one of its neighbors with equal probability and that

the average number of nodes within a region of radius r is proportional to r^2, the probability

density function of the distance r between node x and y is

$$f(r) = 2r, \quad 0 < r < 1$$

where we have normalized r with regard to R by setting $R = 1$.

Now we can calculate P_{ws} as follows:

$$P_{ws} = \int_0^1 2r P_{ws}(r) dr$$

$$= 2p'(1 - p') e^{-p'N} \int_0^1 r e^{-p' \lambda B(r)(2l_{rts}+1)} dr$$

$$= 2p'(1 - p') e^{-p'N} \int_0^1 r e^{-p'N[1 - 2q(r/2)/\pi](2l_{rts}+1)} dr. \qquad (2.3)$$

22

From the Markov chain shown in Figure 2.2, the transition probability P_{ww} that node x continues to stay in *wait* state in a slot is just $(1 - p')e^{-p'N}$, i.e., node x does not initiate any transmission and there is no node around it initiating a transmission. Let π_s, π_w and π_f denote the steady-state probability of state *succeed*, *wait* and *fail*, respectively. From Figure 2.2, we have

$$\pi_w P_{ww} + \pi_s + \pi_f = \pi_w$$

$$\pi_w P_{ww} + 1 - \pi_w = \pi_w$$

$$\pi_w = \frac{1}{2 - P_{ww}} = \frac{1}{2 - (1 - p')e^{-p'N}}.$$

Therefore, the steady-state probability of state *succeed*, π_s, can be calculated as:

$$\pi_s = \pi_w P_{ws} = \frac{P_{ws}}{2 - (1 - p')e^{-p'N}} = p_s. \tag{2.4}$$

Equation (2.4) points out the fact that π_s is just the previous unknown quantity p_s in Equation (2.1). Combining Equations (2.1), (2.3) and (2.4) together, we get a complex relationship between p and p'. However, given p, p' can be computed easily with numerical methods.

Accordingly, the throughput Th is:

$$Th = \frac{\pi_s \cdot l_{data}}{\pi_w T_{wait} + \pi_s T_{success} + \pi_f T_{fail}}$$

$$= \frac{l_{data} \pi_s}{\tau \pi_w + (l_{rts} + l_{cts} + 2\tau)(1 - \pi_w - \pi_s) + (l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau)\pi_s}$$

$$= \frac{l_{data} P_{ws}}{\tau + (l_{rts} + l_{cts} + 2\tau)(1 - P_{ws} - (1 - p')e^{-p'N}) + (l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau)P_{ws}}.$$

$$\tag{2.5}$$

From the formula used to calculate throughput, we can see that π_s and π_w, from which throughput is derived, are largely dependent on p' and not on p, which is the basis for our simplification of the modeling of the channel presented earlier.

23

To apply our analysis to MAC protocols in which perfect collision avoidance is not enforced, e.g., the IEEE 802.11 MAC protocol, we propose a simple though not rigorous extension of the analysis. We can add another state to the Markov chain for the node model (ref. Figure 2.2) whose duration is $l_{rts} + l_{cts} + l_{data} + 3\tau$. This is a *pseudo-succeed* state in which an RTS-CTS-data handshake takes place without acknowledgment coming back due to collisions, i.e., it is a state derived from the *succeed* state of the perfect collision avoidance protocol. We use an "imperfectness factor" β to model the deviatory behavior of the protocol, given that different MAC protocols may have different values of β. The transition probability from *wait* to the *pseudo-succeed* state is then βP_{ws}, and the transition probability from *wait* to *succeed* is $(1 - \beta)P_{ws}$. Hence, the modified formula for throughput is simply:

$$Th = (1 - \beta)l_{data}\pi_s[\tau\pi_w + (l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau)(1 - \beta)\pi_s$$

$$+ (l_{rts} + l_{cts} + 2\tau)(1 - \pi_w - \pi_s) + (l_{rts} + l_{cts} + l_{data} + 3\tau)\beta\pi_s]^{-1}$$

$$= (1 - \beta)l_{data}P_{ws}[\tau + (l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau)(1 - \beta)P_{ws}$$

$$+ (l_{rts} + l_{cts} + 2\tau)(1 - P_{ws} - (1 - p')e^{-p'N}) + (l_{rts} + l_{cts} + l_{data} + 3\tau)\beta P_{ws}]^{-1}.$$

(2.6)

When the deviatory factor β equals zero, Equation (2.6) is reduced to Equation (2.5).

2.3   Numerical Results

In this section, we compare the throughput of the RTS/CTS scheme with a non-persistent CSMA protocol in which there is a separate channel over which acknowledgments are sent in zero time and without collisions. The performance of the latter protocol in multi-

hop networks has been analyzed by Wu and Varshney [72] and we should note that, in practice, the performance of the CSMA protocol would be worse as both data packets and acknowledgments are transmitted in the same channel.

We present results when either relatively large data packets or relatively small data packets are sent. Let τ denote the duration of one time slot. RTS, CTS and ACK packets last 5τ. As to the size of data packets, we consider two cases. One case corresponds to a data packet that is much larger than the aggregate size of RTS, CTS and ACK packets. The other case corresponds to a data packet being only slightly larger than the aggregate size of RTS, CTS and ACK packets. In the latter case, which models networks in which radios have long turn-around times and data packets are short, it is doubtful whether a collision avoidance scheme should be employed at all, because it represents excessive overhead.

We first calculate throughput with different values of α, which we define as the ratio between the circular region including nodes affected by an RTS/CTS handshake and the largest possible circular region in which nodes are guaranteed to be connected with one another. We find that, though the relationship between the ready probability p and transmission-attempt probability p' under different values of α might be somewhat different, the throughput is largely unaffected by α, which is shown in Figure 2.4.[1] In Figure 2.4, N is the average number of nodes that compete against one another to access the shared channel. Thus, the burden of estimating α is relieved in our model, and we can focus on the case in which $\alpha = 1$ thereafter. However, as a side effect of not knowing the actual α that should be used, the relationship between p' and throughput may not agree with the simulations. However, for our

---

[1] The curves for $N = 3$ with different values of α concentrates on the upper part of these figures while the ones for $N = 10$ on the lower part.

PSfrag replacements
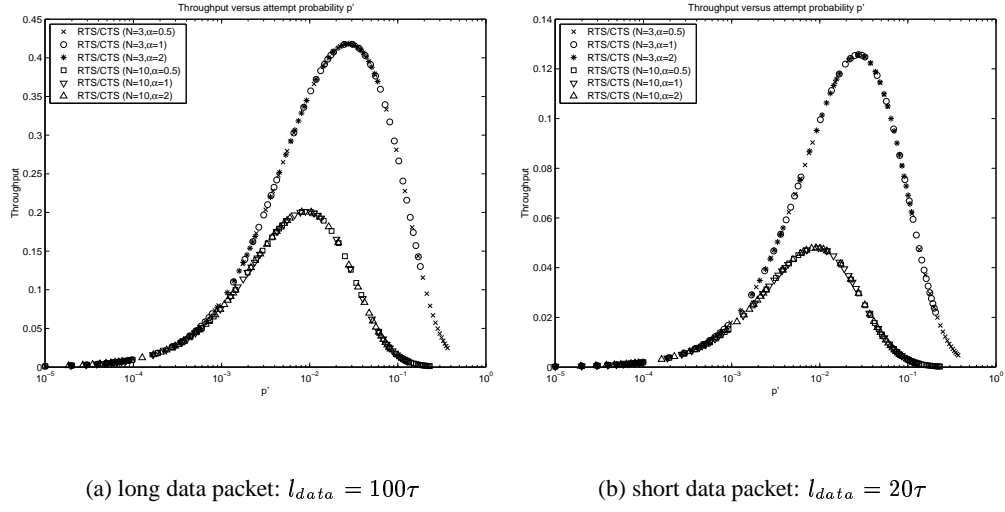
Figure 2.4: α's influence ($l_{rts} = l_{cts} = l_{ack} = 5\tau$)

(a) long data packet: $l_{data} = 100\tau$

(b) short data packet: $l_{data} = 20\tau$
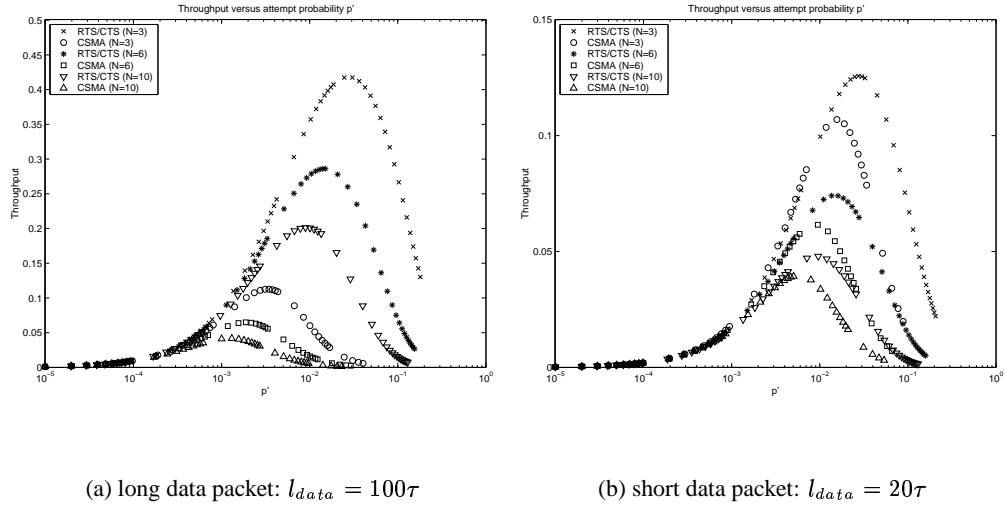
Figure 2.5: Throughput comparison ($l_{rts} = l_{cts} = l_{ack} = 5\tau$)

(a) long data packet: $l_{data} = 100\tau$

(b) short data packet: $l_{data} = 20\tau$

purposes this is not a problem, because we are interested in the saturated throughput only.

Figure 2.5 compares the throughput of collision avoidance against that of CSMA with different values of N and data packet lengths, and we can make the following observations from the above results.

26

When data packet is long, the throughput of CSMA is very low, even for the case in which only $N = 3$ nodes are competing for the shared channel. By comparison, the RTS/CTS scheme can achieve much higher throughput, even when the average number of competing nodes is 10. The reason is simple, the larger a data packet is, the worse the impact of hidden terminals is for that packet in CSMA, because the vulnerability period becomes twice the length of the data packet. With collision avoidance, the vulnerability period of a handshake is independent of the length of data packets, and in the worse case, equals twice the length of an RTS. When a data packet is not very long and the overhead of the collision avoidance and handshake seems to be rather high, collision avoidance can still achieve marginally better throughput than CSMA. We need to emphasize that the performance of the actual CSMA protocol would be much worse than the idealized model we have used for comparison purposes, because of the effect of acknowledgments.

Despite the advantage of collision avoidance, its throughput still degrades rapidly with the increase of N. This is also evident for low values of p' as shown in Figure 2.5. This is due the fact that nodes are spending much more time on collision avoidance and backoff. When N increases, p' decreases much slower to achieve optimum throughput, which already decreases. This shows that collision avoidance becomes more and more ineffective when the number of competing nodes within a region increases, even though these nodes are quite "polite" in their access to the shared channel. This is also different from a fully-connected network, in which the maximum throughput is largely indifferent to the number of nodes within a region [32].

Our results also reveal that hidden terminals degrade the performance of collision

avoidance protocols beyond the basic effect of having a longer vulnerability period for RTSs. There is one dilemma here. On the one hand, it is very difficult to get all the competing nodes around one node coordinated well by probabilistic methods such as randomized backoff. Here the competing nodes refer to both one-hop and two-hop neighbors[2] of the node. In actual MAC protocols, the collisions of data packets may still occur and throughput degrades with increasing numbers of neighbors. On the other hand, even if all the competing nodes of one node defer their access for the node, the possible spatial reuse in multi-hop networks is greatly reduced and hence the maximum achievable throughput is reduced. This dilemma leads to the scalability problem of contention-based MAC protocols that occurs much earlier than people might expect, as the throughput is already quite meager when the average of competing nodes within a region (N) is only 10.

2.4 Simulation Results

The numerical results in the previous section show that an RTS/CTS based access scheme outperforms CSMA, even when the overhead of RTS/CTS packets is comparable to the data packets to be transmitted if perfect collision avoidance can be achieved. In this section, we investigate the performance of the popular IEEE 802.11 DFWMAC protocol, and one of its variants that uses fixed contention window, to validate the predictions made in the analysis.

We use GloMoSim 2.0 [73] as the network simulator. Direct sequence spread spectrum (DSSS) parameters are used throughout the simulations, which are shown in Table 2.1.

---

[2]Here we refer to those nodes that have at least one common neighbor with a node but are not direct neighbors of the node as the node's two-hop neighbors.
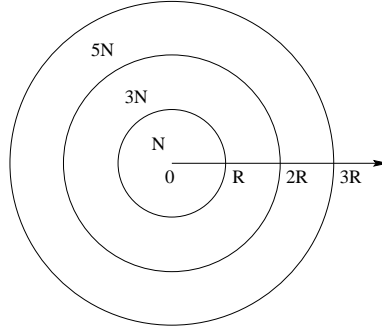
Figure 2.6: Network model illustration

The raw channel bit rate is 2Mbps. We use a uniform distribution to approximate the Poisson distribution used in our analytical model, because the latter is mainly used to facilitate our derivation of analytical results. In addition, it is simply impractical to generate 2, 3, 4, ... nodes within one region, get the throughput for the individual configuration and then calculate the average like what is required in the analytical model. In the network model used simulations, we place nodes in concentric circles or rings as illustrated in Figure 2.6. That is, given that a node's transmitting and receiving range is R and that there are on average N nodes within this circular region, we place N nodes in a circle of radius R, subject to a uniform distribution. Because there are on average $2^2 N$ nodes within a circle of radius $2R$, we place $2^2 N - N = 3N$ nodes outside the previous circle of radius R but inside the concentric circle of radius $2R$, i.e., the ring with radii R and $2R$, subject to the same uniform distribution. Then $3^2 N - 2^2 N = 5N$ nodes can be placed in an outer ring with radii $2R$ and $3R$.

Because it is impossible to generate the infinite network we assumed in our analysis in simulations, we just focus our attention on the performance of the innermost N nodes. Another reason is that it is more appropriate to investigate the performance of MAC schemes in a local neighborhood, rather than in the whole network, because totaling and averaging

performance metrics such as throughput and delay with regard to all the nodes both in the center and at the edge of a network may lead to some askew results. For example, nodes at the edge may have exceedingly high throughput due to much less contention and including them in the calculation would lead to higher than usual throughput. In our experiments, we find that nodes that are outside the concentric circles of radius $3R$ almost have no influence on the throughput of the innermost N nodes, i.e., boundary effects can be safely ignored when the circular network's radius is $3R$. Accordingly, we present only the results for a circular network of radius $3R$.

The backoff timer in the IEEE 802.11 MAC protocol is drawn from a uniform distribution whose upper bound varies according to the estimated contention level, i.e., a modified binary exponential backoff. Thus, p' takes on dynamic values rather than what we have assumed in the analytical model. Accordingly, we expect that the IEEE 802.11 MAC protocol will operate in a region, while our analysis gives only average performance. In addition, even in network topologies that satisfy the same uniform distribution, we can still get quite different results, which will be shown later.

As we have stated, the IEEE 802.11 MAC protocol cannot ensure collision-free transmission of data packets, even under the assumption of perfect carrier sensing and collision avoidance. There are two reasons for this. One is that the length of a CTS is shorter than that of an RTS, which has been shown to prevent some hidden nodes from backing off [17]. The other reason is that, when a node senses carrier in its surroundings, it does not defer access to the channel for a definite time (which is implicit in other protocols [17]) after the channel is clear. When the interfering node perceives the channel idle and a packet from the upper

30

(a) B receives noise.

(b) B receives noise again, but is not forced to defer.

(c) B has a packet in buffer and sends RTS which collides with ACK from D.
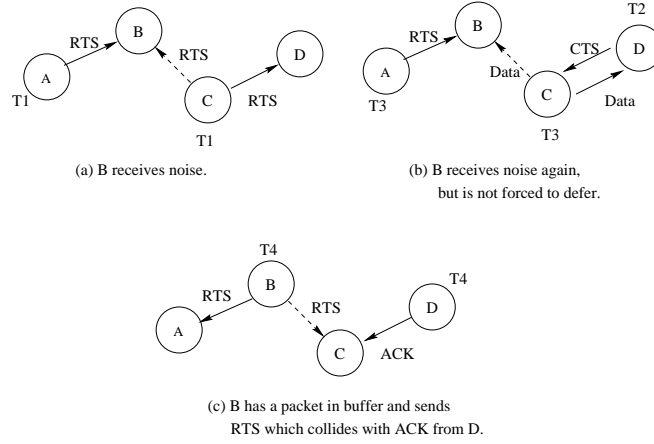
Figure 2.7: Example of collisions with data packets in the IEEE 802.11 MAC protocol

layer happens to arrive in its buffer, it may transmit immediately after the channel is idle for a DIFS (Distributed InterFrame Space) time, while in fact a data packet transmission may still be going on between another two nodes and collision will occur! This can be illustrated by the simple example shown in Figure 2.7.

In our simulation, each node has a constant-bit-rate (CBR) traffic generator with data packet size of 1460 bytes, and one of its neighbors is randomly chosen as the destination for each packet generated. All nodes are always backlogged. Considering the physical layer's synchronization time as well as propagation delay used in the simulation, the effective packet transmission times are shown in Table 2.1. For comparison purposes, we map these simulational parameters to equivalent parameters in our analytical model and they are shown in Table 2.2.

We run both analytical and simulation programs with $N = 3$, 5 and 8. Though we have not tried to characterize how the performance of the IEEE 802.11 MAC protocol is distributed in the region of values taken by p', we do have generated 50 random topologies that

Table 2.1: IEEE 802.11 protocol configuration parameters

RTS	CTS	data	ACK	DIFS	SIFS
20-byte	14-byte	1460-byte	14-byte	50μsec	10μsec
contention window		slot time	sync. time		prop. delay
31–1023		20μsec	192μsec		1μsec

Table 2.2: Equivalent configuration parameters for the analytical model

	τ	l_{rts}	l_{cts}, l_{ack}	l_{data}
actual time	21μsec	272μsec	248μsec	6032μsec
normalized	1	13	12	287

satisfy the uniform distribution and then get an average transmission probability and throughput for the N nodes in the innermost circle of radius R for each configuration. The results are shown in Figure 2.8, in which the centers of rectangles are the mean values of p' and throughput and their half widths and half heights are the variance of p' and throughput, respectively. These rectangles roughly describe the operating regions of IEEE 802.11 MAC protocol with the configurations we are using.

Figure 2.8 clearly shows that, IEEE 802.11 cannot achieve the performance predicted in the analysis of correct collision avoidance, but may well outperform the analysis with the same p' for some configurations, especially when N is small. On first thought, it may seem contrary to intuition, given that IEEE 802.11 cannot ensure collision-free data packet transmissions and should always perform worse than analysis results. In fact, the exceedingly high throughput is largely due to the unfairness of the binary exponential backoff (BEB) used in IEEE 802.11. In BEB, a node that just succeeds in sending a data packet resets its contention window to the minimum value, through which it may gain access to the channel again
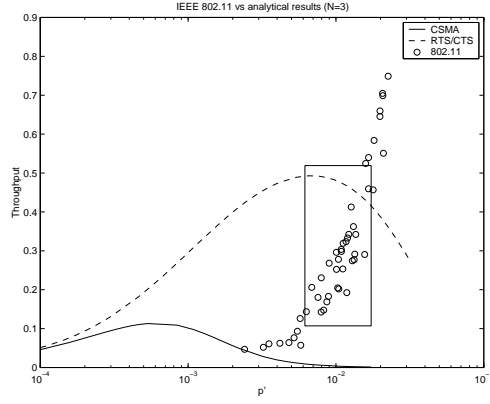
(a) $N = 3$
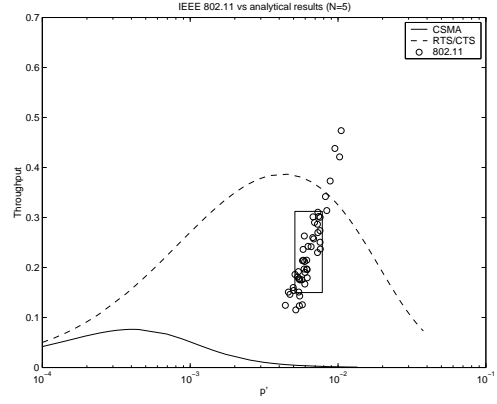
(b) $N = 5$
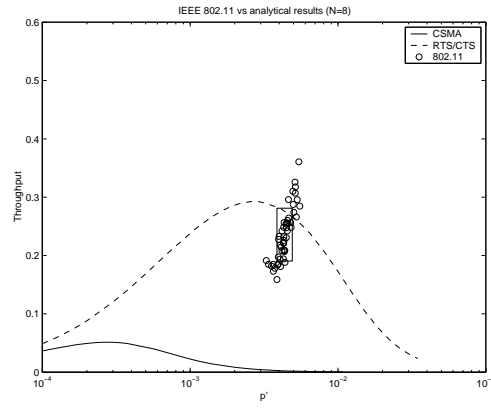
(c) $N = 8$

Figure 2.8: Performance comparison of IEEE 802.11 with analytical results

much earlier than other surrounding nodes. Thus, a node may monopolize the channel for a very long time during which there is no contention loss and throughput can be very high for a particular node, while other nodes suffer starvation. We also find that when N increases, the variance of p' and throughput becomes smaller. Thus, the fairness problem is less severe when there are more nodes competing in a shared channel.

Due to the inherent deficiency of the BEB scheme used in the IEEE 802.11 MAC protocol, we investigate a simple variant of this protocol in which the contention window (*CW*) is fixed and then the backoff timer is generated from a uniform distribution with values ranging between 0 and *CW*. We vary the *CW* to get different values of p' and throughput. Though this modified protocol is not fine tuned to the actual number of neighbors that a node may have and thus is not able to deliver the best performance, it is still a much fairer scheme that reflects more realistically how well a contention-based MAC protocol may perform. In order to have a fair comparison of this scheme with the original IEEE 802.11, we reuse the aforementioned network configurations. The *CW* used in our simulations are tabulated in Table 2.3. The simulation results are shown in Figure 2.9. For clarity, Figure 2.9 shows only the operating regions (shown in rectangles) of the modified IEEE 802.11 protocol, without showing details of how each set of the 50 configurations performs. In addition, the median values of p' and throughput are drawn in Figure 2.9 to show how the throughput is affected by p' or *CW*, where a larger *CW* means a smaller p'.

Comparing Figures 2.8 and 2.9, it is very clear that the modified IEEE 802.11 protocol with a fixed *CW* has a smaller variance of throughput than that in the original protocol and thus is much fairer. To demonstrate this point quantitatively, we obtain both the maximum and

Table 2.3: Contention window (CW) used in simulations

	CW1	CW2	CW3	CW4	CW5	CW6	CW7	CW8
N=3	20	40	80	160	320	640	1280	2560
N=5	30	60	120	240	480	960	1920	3840
N=8	50	100	200	400	800	1600	3200	6400

(a) $N = 3$

(b) $N = 5$

(c) $N = 8$
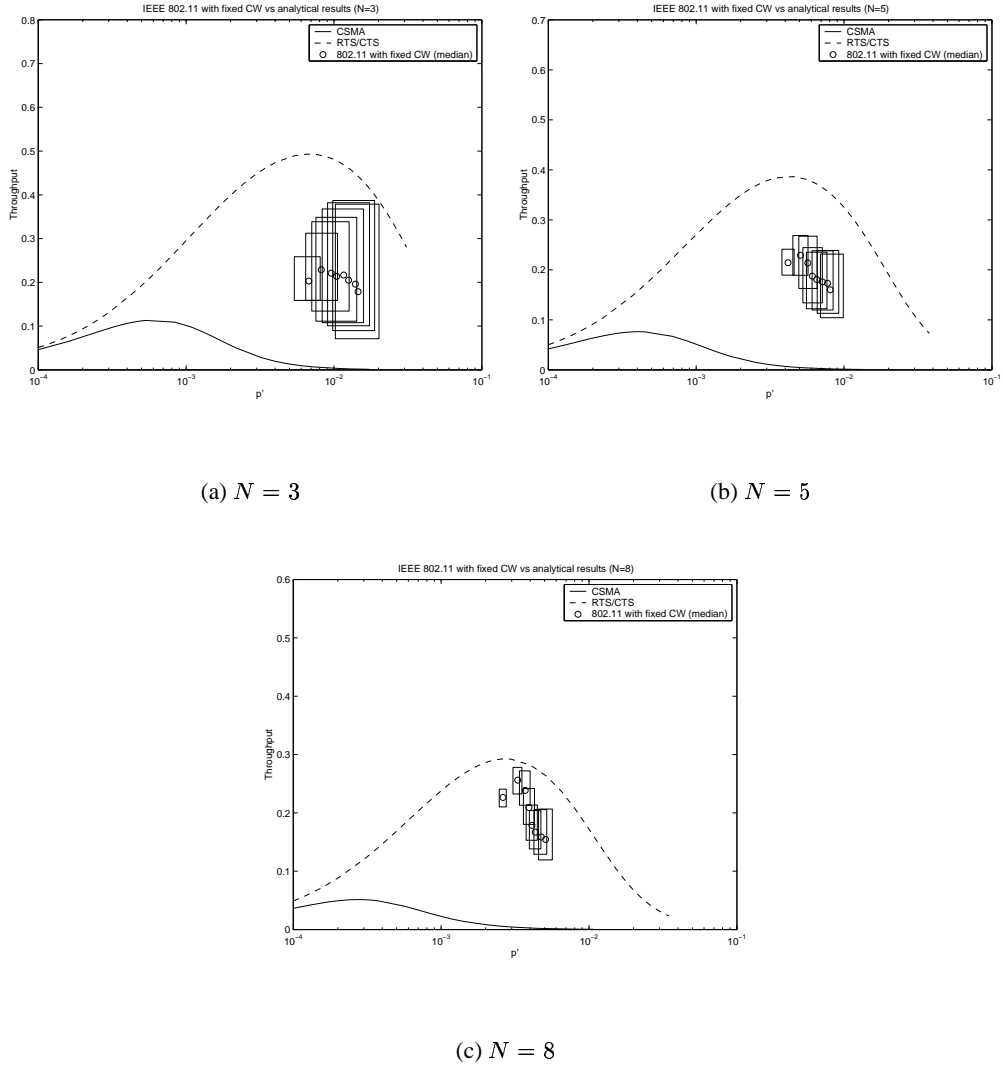
Figure 2.9: Performance comparison of IEEE 802.11 (fixed CW) with analytical results

Table 2.4: Fairness comparison of BEB scheme and fixed *CW* scheme

	original	CW1	CW2	CW3	CW4	CW5	CW6	CW7	CW8
N=3, mean	10.51	3.77	3.25	3.00	2.99	3.34	2.28	1.81	1.46
N=3, std	15.67	5.30	3.02	2.52	2.90	4.65	1.35	0.75	0.39
N=5, mean	4.54	3.42	2.89	2.65	2.46	2.29	2.04	1.73	1.46
N=5, std	2.99	1.83	1.21	0.89	0.80	0.60	0.53	0.36	0.22
N=8, mean	3.45	3.57	2.88	2.75	2.39	2.16	1.95	1.69	1.46
N=8, std	1.59	3.11	1.47	1.70	0.87	0.73	0.58	0.35	0.39

the minimum throughput among the innermost N nodes and then calculate the ratio between the maximum and the minimum as an index for fairness. The smaller the ratio is, the fairer is the protocol and vice versa. The results are shown in Table 2.4. In addition, we can see the degraded performance in the fixed *CW* scheme due to more contention, especially when N is small.

Given that these two protocols cannot ensure that data packets are transmitted free of collisions, the throughput can deviate substantially from what is predicted in the analysis. To demonstrate this, we also collect statistics about the number of transmitted RTS packets that will lead to ACK timeout due to collision of data packets as well as the total number of transmitted RTS packets that can lead to either an incomplete RTS-CTS-data handshake or a successful four-way handshake. Then we calculate the ratio of these two numbers and tabulate the results in Table 2.5.

Table 2.5 clearly shows that much of the precious channel resource is wasted in sending data packets that cannot be successfully delivered. In addition, in order to decrease the percentage of channel resource wasted due to collisions, a larger contention window should be chosen to artificially decrease the transmission probability of nodes which at the same

36

Table 2.5: Percentage of ACK timeout in BEB scheme and fixed *CW* scheme

	original	CW1	CW2	CW3	CW4	CW5	CW6	CW7	CW8
N=3, mean	0.29	0.49	0.46	0.43	0.41	0.39	0.33	0.23	0.14
N=3, std	0.17	0.20	0.18	0.17	0.17	0.17	0.15	0.11	0.07
N=5, mean	0.39	0.56	0.54	0.53	0.51	0.47	0.37	0.25	0.15
N=5, std	0.10	0.10	0.10	0.10	0.10	0.10	0.09	0.07	0.04
N=8, mean	0.44	0.59	0.58	0.56	0.53	0.44	0.33	0.21	0.14
N=8, std	0.06	0.07	0.07	0.06	0.06	0.06	0.05	0.04	0.07

time leads to longer time wasted in waiting. This is a very typical behavior of collision-avoidance protocols, especially those protocols that do not have a correct collision avoidance scheme. The possibility of collisions of data packets with other packets places a limit on the maximum achievable throughput, which can be significantly lower than the theoretical results that assume a perfect collision avoidance.

Figures 2.8 and 2.9 show that the gap in maximum throughput between analytical and simulation results decreases when N increases. This can be explained as follows. When the number of direct competing nodes N increases, the number of indirect competing nodes (hidden terminals, $3N$ on average) also increases, which makes nodes implementing a perfect collision avoidance protocol spend much more time in deferring and backing off to co-ordinate with both direct and indirect competing nodes to avoid collisions. Therefore, much of the gain of perfect collision avoidance is lost and possible spatial reuse is also reduced in a congested area, which makes a perfect collision avoidance protocol work only marginally better than an imperfect one. This observation could not be predicted from previous analytical models or simulations focusing on fully-connected networks or networks with only a limited number of hidden terminals [32–34].

(a) $N = 3$ (b) $N = 5$
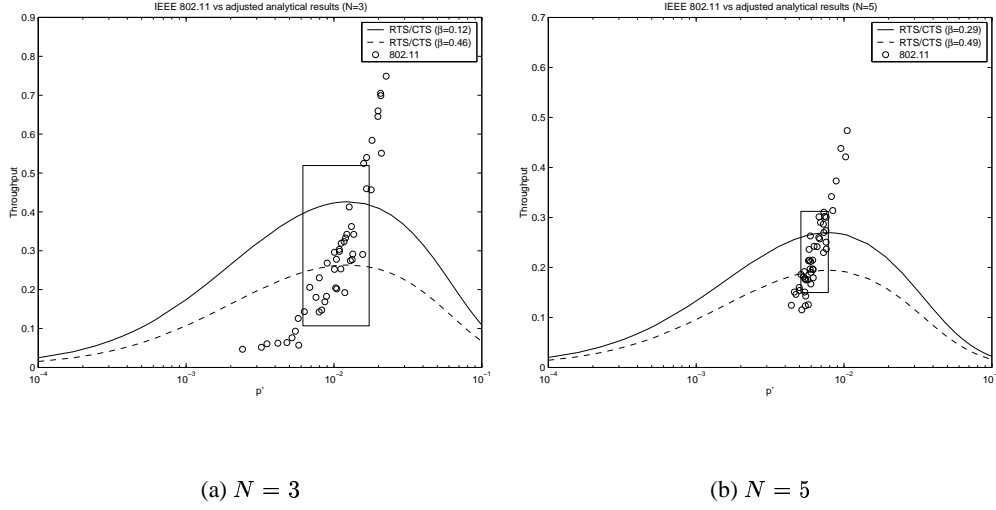
Figure 2.10: Performance comparison of IEEE 802.11 with adjusted analytical results

The percentage shown in Table 2.5 is in fact the β in our extended analysis to explain the deviatory behavior of MAC protocols that do not have perfect collision avoidance. Using these values, we compare the performance of the IEEE 802.11 protocol with that of the adjusted analysis obtained from Equation (2.6), and show the results in Figure 2.10. In Figure 2.10, we only show the results for small values of N as it is not quite meaningful to do the adjustment for large values of N due the reason stated above. Figure 2.10 shows that the extended analysis is a rather good approximation of the actual performance of the IEEE 802.11 protocol though the latter has larger variation in throughput (possibly due to its inherent fairness problems).

2.5 Conclusion

In this chapter, we have used a simple model to derive the saturation throughput of MAC protocols based on an RTS-CTS-data-ACK handshake in multi-hop networks. The results show that these protocols outperform CSMA protocols, even when the overhead of RTS/CTS exchange is rather high, thus showing the importance of correct collision avoidance in random access protocols. More importantly, it is shown that the overall performance of the sender-initiated collision avoidance scheme degrades rather rapidly when the number of competing nodes allowed within a region increases, in contrast to the case of fully-connected networks and networks with limited hidden terminals reported in the literature [32–34], where throughput remains almost the same for a large number of nodes. The significance of the analysis is that the scalability problem of contention-based collision-avoidance MAC protocols looms much earlier than people might expect. Simulation experiments with the IEEE 802.11 MAC protocol and one of its variants validate these observations and show that the IEEE 802.11 MAC protocol can suffer severe degradation in throughput due to its inability to avoid collisions between data packets and other packets even when the number of competing nodes in a region is small. However, when the number of competing nodes in a region increases, the performance gap is smaller as perfect collision avoidance protocols also begins to suffer from exceedingly long waiting time.

Based on both analytical and simulation results, we observe that there are some possible ways to improve the throughput of a sender-initiated collision-avoidance protocol in ad hoc networks. First of all, the simulation results show that it is very important to ensure correct collision avoidance when the network is less crowded. Using the longer CTS packets

proposed in [22], the IEEE 802.11 MAC protocol can lead to much better performance in throughput.

Another obvious way to improve the performance of the IEEE 802.11 protocol is to reduce τ, which includes carrier sensing delay, transmit to receive turnaround time, etc., so as to enlarge the ratio of data packet transmission time to τ. In effect, this implies reducing the transmission power of the nodes and reducing the length of control overhead. Given that RTS and CTS packets cannot be reduced in length and arguably the CTS needs to be lengthened to be sent as a busy tone, the latter requires using piggyback acknowledgments or making acknowledgments optional. An alternative is to combine the sender-initiated handshake with receiver-initiated handshake, as the latter is shown to have less control overhead [24]. However, receive-initiated schemes generally require a good traffic estimator and an adaptive polling discipline at polling nodes to work well, which have not been investigated thoroughly so far.

Because the optimum value of p', i.e., the probability that a node transmits in a time slot, changes with the number of competing nodes within a region, it is necessary to have an adaptive algorithm to achieve optimum performance when the number of active nodes within a region changes. For example, Cali et al. [34] have proposed a way to dynamically tune the IEEE 802.11 protocol to achieve better performance, though it is investigated in static networks. Given that the original BEB scheme has inherent fairness problem and the fixed contention window does not adapt well, it is fair to say that there is still much work left to be done on this topic. We will address the fairness problem in detail in Chapters 5, 6 and 7.

Recently, MAC protocols using directional antennas have also been proposed [35–

37] and shown to perform much better than the existing omni-directional MAC protocols. However, most of the work is simulation based and the lack of analytical modeling motivates us to evaluate the performance of directional MAC protocols via both analytical and simulation methods. We will present our work in Chapters 3 and 4.

Chapter 3

Directional MAC Schemes with

Unicast Traffic

3.1   Introduction

The four-way collision avoidance scheme used in the original IEEE 802.11 MAC protocol depends on the broadcast nature of the channel and assumes that all nodes are equipped with omni-directional antennas. However, this scheme requires that all the neighbors of the sending and receiving nodes back off during the handshake which greatly reduces the possible spatial reuse in multi-hop networks.

Based on the above observations, some MAC protocols that make use of directional antennas have been proposed and studied in the recent past [35–41], which we call *directional collision avoidance* protocols. Ko et al. [35] propose two schemes. One scheme consists of nodes using directional RTS transmissions and omni-directional transmission of CTS packets

42

in collision avoidance, and then using directional transmissions of data and acknowledgment packets after successful handshakes. The other scheme consists of nodes using both directional and omni-directional transmission of RTS packets alternately. When the location of a receiver is not well known or all the transmitting antennas are unblocked, an omni-directional RTS is sent. These two schemes show the tradeoff between increased possibility of simultaneous transmissions by neighboring nodes (scheme one) and reduced possibility of collisions of control packets (scheme two).

Nasipuri et al. [36] propose a MAC protocol similar to those summarized above, but use a different model. In the authors' model, each node is equipped with M antennas whose orientations can be maintained all the time irrespective of a node's movement. They also assume that nodes have directional reception capability, that is, that they can activate the antenna pointing to the desired source while deactivating antennas in other directions. Therefore, the receiving node is not influenced by simultaneous transmissions from other directions. This is different from the model assumed by Ko et al. [35] where antennas are always active for receiving and thus transmissions to different antennas result in failed reception. In the proposed MAC protocol, omni-directional RTS and CTS packets are first exchanged between a pair of sending and receiving nodes and then directional transmissions of data and acknowledgment packets are used.

Ramanathan [37] studies the performance of beamforming antennas in ad hoc networks when RTS and CTS packets are transmitted omni-directionally with the same range as directional transmissions while data and ACK packets are transmitted directionally. The author also addresses some interesting problems arising from directional transmissions, such

as link power control and directional neighbor discovery.

In the work done by Takai et al. [38] and Choudhury et al. [39], direction information is included in the network allocation vector (NAV), which is used by nodes in IEEE 802.11 to notify other nodes how long they should defer their access to the shared channel. With detailed direction information, a node receiving NAV from one direction can still transmit in other directions. Choudhury et al. [39] also propose a multi-hop RTS scheme to take advantage of the higher gain in directional transmissions. The authors also show the importance of considering the different ranges of omni-directional and directional transmissions as the results may be quite different from those when omni-directional and directional transmissions are assumed to have similar ranges.

Simulation studies of the above proposed protocols show that they improve performance over the existing omni-directional IEEE 802.11 MAC protocol. However, the majority of the performance analyses of directional collision avoidance schemes have been done via simulations [35–39], and there is little prior work on the analytical modeling of directional collision avoidance protocols. In this chapter, we investigate the interaction between spatial reuse, interference reduction and collision avoidance by extending the simple yet general network model presented earlier in Chapter 2 and derive the analytical results. An important contribution of this model is that it is applicable to many other combinations of directional and omni-directional transmissions in collision avoidance protocols.

Section 3.2 outlines the directional collision avoidance schemes that we study analytically and by simulation. The schemes we address consider that nodes communicate directly only with neighboring nodes within their omni-directional transmission range even

44

when directional transmissions may have higher gains and reach more nodes. The rationale for this approach is that, obtaining information about the relative location of neighboring nodes can be easily done with broadcast transmissions, thus avoiding the complexity of discovering neighbors who can only be reached by directional transmissions.

Section 3.3 presents the first analytical study of directional collision avoidance in ad hoc networks that considers (a) the effect of directional transmitting and receiving on spatial reuse and collision avoidance, and (b) the effect of the differences in gains between omni-directional and directional transmissions. To attain a tractable analytical model, we assume that interference due to side lobes outside the transmission beamwidth are negligible. Although this is not the case with real directional antennas, our model does provide a good approximation for ad hoc networks in which any node communicates with only those nodes that reside within its omni-directional transmission range. The results of the analysis show that the scheme that uses both narrow-beam directional transmissions and receptions throughout the collision-avoidance handshake can achieve the best performance among all the schemes investigated, and that one-hop throughput does not degrade when the number of competing nodes within a neighborhood increases because of the much increased spatial reuse in narrow beamwidth transmission/reception. It is also shown that, as expected, the performance of directional collision avoidance schemes degrades when directional transmissions have much higher gain than omni-directional transmissions, because of the increased interference range and reduced spatial reuse.

Section 3.4 presents the results of simulations carried out to validate the results from the analytical model, as well as to analyze the effect of side lobes in directional transmissions.

The IEEE 802.11 MAC protocol and its variants that implement directional collision avoidance are investigated.

The results obtained via simulations validate the results predicted by our analytical model. It is further shown that directional reception helps to cancel out almost all the adverse effects of hidden terminals, and achieves very low data packet collision ratio. The simulation results also show that the effects on side lobes on throughput are negligible if the gain of the main transmission lobe is reasonably higher than that of side lobes and the carrier sensing threshold is raised to make nodes less sensitive to channel activities.

Section 3.5 summarizes our results and outlines directions for future work.

3.2   Directional Collision Avoidance Schemes

In this section, we classify the schemes according to whether omni-directional or directional transmitting/receiving is used.

In the omni-directional MAC scheme, all packet transmissions and receptions are omni-directional. The IEEE 802.11 MAC protocol is an example of a protocol based on this scheme. We call this scheme *OTOR* (for "omni-directional transmission and omni-directional reception"). In our analysis, we assume that correct collision avoidance is enforced, i.e., once a node starts sending a CTS packet in reply to an RTS destined to it, the following handshake can go on unobstructed.[1] This scheme has been analyzed thoroughly in Chapter 2 and serves as the basis for comparison.

For directional collision avoidance schemes, we first consider two directional collision-

---

[1] The reader is referred to [22] for a discussion of issues involved in achieving correct collision avoidance.

avoidance schemes in our analytical study. In both schemes, RTS, CTS, data packets and ACK are transmitted directionally. When a node is transmitting in one direction, it appears "deaf" to other directions and cannot sense any channel activity at all. In these schemes, spatial reuse is maximized as nodes limit their transmission to as small area as possible. They differ in whether directional receiving is used or not.

One scheme uses omni-directional receiving mode whenever it is not transmitting; we call such a scheme *DTOR*, which is short for "directional transmission and omni-directional reception." In the other scheme, which we call *DTDR* (for "directional transmission and directional reception"), a node directs its antenna to the neighbor from which it expects to receive a packet, such as a CTS, a data packet, or an ACK, and appears "deaf" to transmissions from other directions. A node that is not expecting a packet from a neighbor listens to the channel omni-directionally. If a node does not receive an expected packet within its due time, the node returns to omni-directional receiving mode.

It is also possible to establish schemes that combine both omni-directional and directional transmissions and receptions. For example, in one scheme, only CTS packets are transmitted omni-directionally, while all the other types of packets are transmitted directionally. The purpose of the omni-directional transmission of CTS packets would be to try to silence the neighbors of a receiver and to prepare a clear *floor* for the sending node. Omni-directional or directional reception can be applied to this approach. When the former is used, we call it *MTOR* scheme (short for "mixed-directional transmission and omni-directional reception"); when the latter is used, we call it *MTDR* scheme (short for "mixed-directional transmission and directional reception").

To implement any of the directional collision avoidance schemes, nodes have to know the relative locations of their neighbors. In practice, exact locations of nodes' neighbors are not required for directional collision avoidance schemes to function. Generally speaking, the angle of arrival (AoA) information reported by the radio can be cached and used later to direct antenna orientation.

In this chapter, we assume that broadcast beacons are used by nodes to determine who their neighbors are and their relative locations. Hence, even though a node can transmit and receive directionally, it communicates directly only with nodes within its omni-directional transmission range. The rationale for this approach derives from the results obtained in our analysis. Our analytical model studies the case in which location information is already known, and computer simulations are used to verify the results from this model and to show that even a crude and simple beaconing mechanism in which nodes let their neighbors know about their existence (and relative location) using periodic broadcast transmissions suffices to support directional collision avoidance schemes without a significant impact on performance.

3.3 Approximate Analysis

In this section, we present the analysis of the following MAC schemes: *DTOR*, *MTOR* and *DTDR*. We then compare their performance with the existing omni-directional *OTOR* scheme. The *MTDR* scheme can be analyzed using the approach presented here, but its analysis is omitted for brevity. However, it is still compared with the other schemes later in simulations.

3.3.1 Assumptions

Most of the key assumptions have been presented in Section 2.2. Here we highlight some of the additional assumptions necessary for extending the analysis to MAC schemes using directional antennas.

Similar to the analysis done in Section 2.2, we assume that a node becomes ready independently with probability p_0 (p in Section 2.2) at each time slot, that a node begins transmission with probability p (p' in Section 2.2) at each time slot and that a node initiates a successful handshake with any other node with probability p_s. In Section 2.2, a Markov chain of the shared channel is used to derive the rough relationship between p and p_0, and a Markov chain is used to model the state of a node to derive the throughput. It is shown that the throughput is largely decided by p.

Due to the workings of collision avoidance and resolution, p must be kept very small, which means that p is likely to assume values that do not exceed 0.1. Here we do not analyze the relationship between p and p_0, as has been done before [72] and Chapter 2. This is because of the difficulty in modeling the channel when the directional transmission mode is used, especially in the mixed-transmission schemes we investigate, in which nodes can switch between directional and omni-directional transmission modes. Instead, we just assume that p takes on a range of values and then derive the throughput using the node model only. Given that the key objective of the model is to provide a comparative analysis of directional collision-avoidance strategies, and that the probability of successful handshakes by any one node in an ad hoc network cannot be very large, our approximation is very sensible.

In our analytical modeling, we also assume complete signal attenuation outside
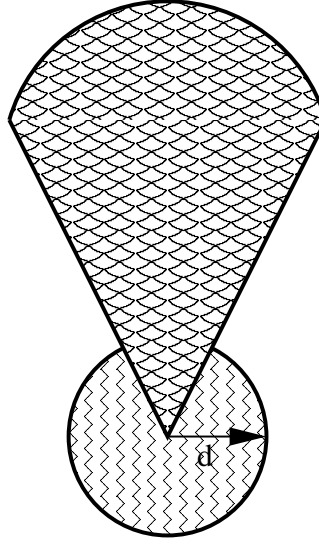
Figure 3.1: Antenna model with side lobes

the main transmission beamwidth. In reality, beamforming antennas can generate side lobes which can cause additional interferences to neighbor nodes which lie outside the main lobe but inside the side lobes. Fig. 3.1 shows a more realistic modeling of directional antennas, where the interference range of side lobes is modeled by the parameter d. However, we reason that the influence of side lobes on throughput is insignificant if the following two approaches are adopted. The first approach is to reduce the side lobe level which is the ratio, measured in decibels (dB), of the amplitude at the peak of the main lobe to the amplitude at the peak of a side lobe. This approach is obvious and can be achieved with the advancement of antenna systems. The second approach is to raise the carrier sensing (CS) threshold so that nodes' access to the shared channel is less sensitive to channel interferences and nodes will not waste too much time on deferring/waiting. When these two approaches are used, side lobes do not degrade throughput much. Simulations are used to validate our conjecture and are presented in Section 3.4.

50

As mentioned before, the range of omni-directional transmissions is R, and the range of directional transmissions is $R' = \gamma R$, where $\gamma \geq 1$. Suppose N is the average number of nodes within a circular region of radius R; therefore, we have $N = \lambda \pi R^2$. Similarly we have $N' = \lambda \pi R'^2 = \gamma^2 N$.

When a node transmits directionally, its transmissions can have longer range than its omni-directional transmissions. The effect of this is that a receiver can have more interfering sources than the nodes within an omni-directional transmission range. In our analysis and simulations, we assume that a node communicates directly only with other nodes that are within its omni-directional transmission range R, and communicate only indirectly with nodes outside R and inside its directional transmission range R', even though it can still be an interfering source for these nodes. The rationale for this assumption is twofold. First, in any directional collision avoidance scheme, a node needs to find the location of the nodes around it to direct its packet transmissions and receptions, and a simple way to accomplish this is by means of omni-directional beacons, especially in ad hoc networks in which nodes can be mobile. Second, it is possible to reduce the transmission power of directional transmissions to attain transmission ranges similar to those of omni-directional transmissions [41], in which case omni-directional and directional transmissions have the same range.

For ease of explanation, we also assume that directional transmissions and receptions have equal beamwidth.

With some simplifications, all the schemes we are investigating share the same node model which is a three-state Markov chain shown in Fig. 2.2 and all states have the same meanings as discussed in Section 2.2. These schemes only differ in the duration of certain

states and the transition probabilities among these states.

3.3.2 The *DTOR* Scheme

Following an approach similar to what we have presented for the *OTOR* scheme analyzed in Section 2.2, it is straightforward to show that the values of many quantities for the *DTOR* scheme remain the same as those in the *OTOR* scheme. They are summarized as follows:

$$T_{succeed} = l_{rts} + l_{cts} + l_{data} + l_{ack} + 4$$

$$T_{wait} = 1$$

$$P_{sw} = P_{fw} = 1$$

$$P_{ww} = (1 - p)e^{-p'N'} \quad (p' = p\theta/2\pi)$$

$$\pi_w = 1/(2 - P_{ww})$$

$$\pi_s = P_{ws}/(2 - P_{ww})$$

$$\pi_f = 1 - \pi_w - \pi_s. \tag{3.1}$$

The only unknown quantities are T_{fail} and P_{ws}. Because the *DTOR* scheme cannot prevent interference from neighboring nodes, the handshake between any pair of sending and receiving nodes may be interrupted at almost any time and the failed period can last from $T_1 = l_{rts} + 1$ to $T_2 = l_{rts} + l_{cts} + l_{data} + l_{ack} + 4$.

We assume that the length of the failed period follows a truncated geometric distribution with parameter p, lower bound T_1 and upper bound T_2. Then we take the mean value
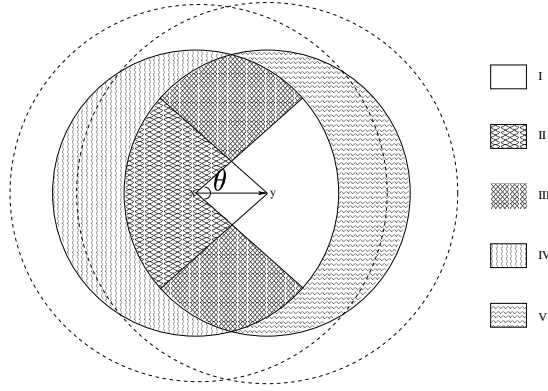
Figure 3.2: The *DTOR* scheme

of the truncated geometric distribution as T_{fail}, which is shown below:

$$T_{fail} = \frac{1-p}{1-p^{T_2-T_1+1}} \sum_{i=0}^{T_2-T_1} p^i (T_1 + i).$$

(3.2)

As discussed in Section 2.2, to calculate P_{ws}, we need to first calculate $P_{ws}(r)$, the

probability of successful transmission in a time slot initiated by node x to node y which is

at a distance r apart. The success of the handshake between nodes x and y depends on the

nodes for which y is within their omni-directional transmission range and those nodes for

which y is within their directional transmission range as illustrated in Fig. 3.2. In Fig. 3.2, θ

is the beamwidth of transmissions and receptions and solid circles indicate omni-directional

transmission ranges of nodes while dashed circles indicate directional transmission ranges.

To simplify our computation of throughput, we assume that there are, in effect, N'

nodes around a node's omni-directional transmission range, though no node is assumed to

communicate directly with any other node that is only reachable from directional transmis-

sions. In fact, this simplifying assumption avoids the complexity of calculating interference

directly from those nodes that are between the solid and dashed circles and instead such inter-

ference is taken into account by increasing the number of nodes within omni-directional range

53

from N to N'. With this simplification, we can proceed as follows.

$P_{ws}(r)$ equals the product of the probability that x transmits, the probability y listens and the probability that no node that can interfere x's transmission to y actually transmits which is denoted by $P_I(r)$, i.e., $P_{ws}(r) = p(1-p)P_I(r)$.

From Fig. 3.2, we can see that the region around nodes x and y can be divided into five areas. Denote by S_i the size of Area i, they are:

$$S_I = \theta/(2\pi)$$

$$S_{II} = \theta/(2\pi) - r^2 \tan(\theta/2)/(2\pi)$$

$$S_{III} = 2q(r/2)/\pi - \theta/\pi + r^2 \tan(\theta/2)/(2\pi)$$

$$S_{IV} = 1 - 2q(r/2)/\pi$$

$$S_V = 1 - 2q(r/2)/\pi \tag{3.3}$$

where we have normalized r with regard to R by setting $R = 1$ and S_i with regard to πR^2. Detailed calculation of these areas is as follows.

For ease of calculation, we set up the coordinate system shown in Fig. 3.3. Given that Area I is a fan, its area is $\theta R^2/2$.

The size of Area II equals the size of the Area I minus the area of the rhombus in the middle. Because line BCD is $y = \tan(\pi - \theta/2)(x - r/2)$ and line ACE is $y = \tan(\theta/2)(x + r/2)$, we can obtain the coordinates of C, which are $(0, r\tan(\theta/2)/2)$. Therefore, the area of the rhombus $ACBF$ is $r^2 \tan(\theta/2)/2$. Accordingly,

$$S_{II} = S_I - S_{ACBF}$$

54
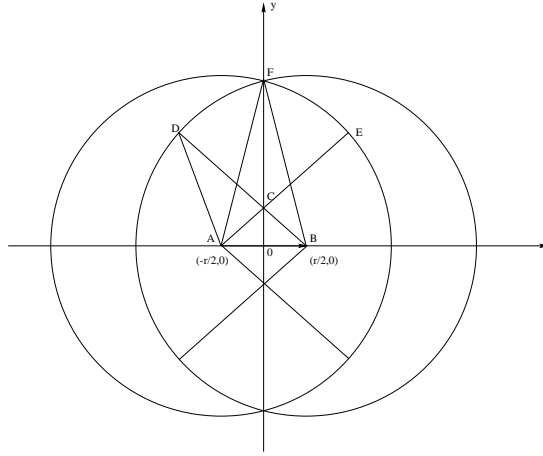
Figure 3.3: Calculation of various areas

$$= \theta R^2/2 - r^2 \tan(\theta/2)/2$$

Given that the total area of I, II and III equals $C(r)$ in [71], Area III can be calculated as follows:

$$S_{III} = C(r) - S_I - S_{II}$$

$$= 2R^2 q(\frac{r}{2R}) - (\theta R^2/2 + \theta R^2/2 - r^2 \tan\frac{\theta}{2}/2)$$

$$= 2R^2 q(\frac{r}{2R}) - \theta R^2 + \frac{r^2}{2} \tan\frac{\theta}{2}.$$

The Areas IV and V are just $B(r)$ derived in [71], which are repeated here:

$$S_{IV} = S_V = \pi R^2 - 2R^2 q(\frac{r}{2R}).$$

Normalizing Areas I–V with respect to πR^2, we obtain (3.3).

Accordingly, for the handshake between nodes x and y to be successful, all of the following conditions should be met:

1. In Area I, no node transmits in one slot and the probability is:

$$p_1 = e^{-pS_I N'} \qquad (3.4)$$

   This is because the nodes in Area I do not know that node x is transmitting and can interfere with the handshake between nodes x and y, if they do transmit.

2. In Area II, it must be true that no node transmits in $2l_{rts}$ slots in the direction of node y and does not transmit in the slot when the transmission of node y arrives at them. Thus the probability is:

$$p_2 = e^{-p' S_{II} N' (2l_{rts})} \cdot e^{-pS_{II} N'} \qquad (3.5)$$

   where $p' = p\theta/(2\pi)$.

3. In Area III, no node transmits in the direction to nodes x and y during the whole hand-shake and the span angle of the direction θ' is $\theta + \phi$, where ϕ is the angle formed by the two lines joining a node in Area III with nodes x and y if ϕ is less than θ; otherwise, θ' is just 2θ. When nodes x and y are very close to each other, $\theta' \approx \theta$. Though the range of θ' is between θ and 2θ, for simplicity, we just choose $\theta' = \theta$. Therefore,

$$p_3 = e^{-p'' S_{III} N' (l_{rts} + l_{rts} + 1 + l_{cts} + 1 + l_{data} + 1 + l_{ack} + 1)}$$

$$= e^{-p'' S_{III} N' (2l_{rts} + l_{cts} + l_{data} + l_{ack} + 4)} \qquad (3.6)$$

   where $p'' = p\theta'/(2\pi) = p\theta/(2\pi)$.

4. In Area IV, no node transmits in node x's direction when node y is transmitting; there-fore, there are two such periods. One is the time when node y transmits a CTS packet

56

to node x and the other is the time when node y transmits an acknowledgment packet

to node x. The durations of these two periods in the number of time slots are approx-

imately $l_{rts} + l_{cts} + 1$ and $l_{rts} + l_{ack} + 1$ respectively which follows the assumption

that nodes transmit in each time slot independently with probability p. Accordingly, the

probability p_4 that no interference from nodes in Area IV is:

$$p_4 = e^{-p'S_{IV}N'(l_{rts}+l_{cts}+1)} \cdot e^{-p'S_{IV}N'(l_{rts}+l_{ack}+1)}$$

$$= e^{-p'S_{IV}N'(l_{rts}+l_{cts}+1+l_{rts}+l_{ack}+1)}$$

$$= e^{-p'S_{IV}N'(2l_{rts}+l_{cts}+l_{ack}+2)} \tag{3.7}$$

5. In Area V, no node transmits in node y's direction when node x is transmitting. Similar

to the case discussed above, there are two such periods. One is the time when node x

transmits an RTS packet to node y and the other is the time when node x transmits a

data packet to node y. The durations of these two periods in the number of time slots are

approximately $l_{rts} + l_{rts} + 1$ and $l_{rts} + l_{data} + 1$ respectively. Therefore, the probability

p_5 that no interference from nodes in Area V is:

$$p_5 = e^{-p'S_V N'(l_{rts}+l_{rts}+1)} \cdot e^{-p'S_V N'(l_{rts}+l_{data}+1)}$$

$$= e^{-p'S_V N'(l_{rts}+l_{rts}+1+l_{rts}+l_{data}+1)}$$

$$= e^{-p'S_V N'(3l_{rts}+l_{data}+2)} \tag{3.8}$$

Having that $P_I(r) = p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot p_5$, P_{ws} can be obtained by considering all

possible values of r, i.e.,

$$P_{ws} = \int_0^1 2r P_{ws}(r) dr.$$

57

Then throughput Th can be calculated as follows:

$$Th = \frac{\pi_s \cdot l_{data}}{\pi_w T_w + \pi_s T_s + \pi_f T_{fail}}$$

$$= l_{data} \pi_s [\pi_w + T_{fail}(1 - \pi_w - \pi_s) + (l_{rts} + l_{cts} + l_{data} + l_{ack} + 4)\pi_s]^{-1}.$$

3.3.3 The *MTOR* Scheme

Once again, following an approach similar to what is used for the *DTOR* scheme, many of the variables needed to compute the throughput of the *MTOR* scheme can be calculated and equal the values shown in (3.1), with the exception that $P_{ww} = (1 - p)e^{-pN'}$, which is product of the probability that the node does not transmit and the probability that no node around it transmits omni-directionally. Though nodes may sometimes transmit directionally, considering that almost each handshake, either failed or successful, consists of a CTS packet that is transmitted omni-directionally, and that the CTS packet can virtually silence all the hidden terminals, this is a reasonable approximation.

To calculate P_{ws}, we first need to calculate $P_{ws}(r)$. In the *MTOR* scheme, the interfering region of nodes x and y can be divided into three areas as shown in Fig. 3.4. Similarly, denote by S_i the size of Area i and they are:

$$S_I = \theta/(2\pi)$$

$$S_{II} = 1 - \theta/(2\pi)$$

$$S_{III} = 1 - 2q(r/2)/\pi.$$
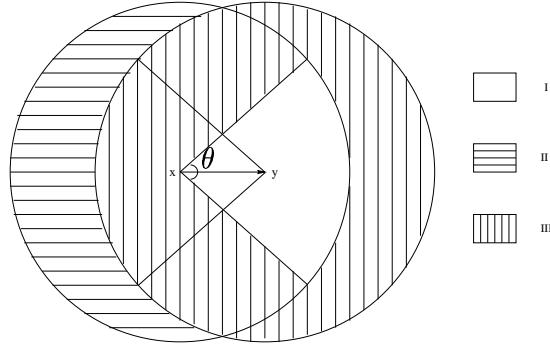
Figure 3.4: The *MTOR* scheme

For the handshake between nodes x and y to be successful, the following conditions should be met:

1. In Area I, no node transmits in one slot and the probability is:

$$p_1 = e^{-pS_I N'}$$

2. In Area II, it should be that no node transmits in $2l_{rts}$ slots in the direction to node y and no node transmits in the slot when the transmission of node y arrives at them. Thus the probability is:

$$p_2 = e^{-p' S_{II} N' (2l_{rts})} \cdot e^{-pS_{II} N'}$$

where $p' = p\theta/(2\pi)$.

3. In Area III, which is in effect Area IV in Fig. 3.2, the probability p_3 that no interference from nodes in the area is:

$$p_3 = e^{-p' S_{III} N' (l_{rts} + l_{cts} + 1)} \cdot e^{-p' S_{III} N' (l_{rts} + l_{ack} + 1)}$$

$$= e^{-p' S_{III} N' (l_{rts} + l_{cts} + 1 + l_{rts} + l_{ack} + 1)}$$

$$= e^{-p' S_{III} N' (2l_{rts} + l_{cts} + l_{ack} + 2)}$$

59

From the above, we can derive $P_{ws}(r)$:

$$P_{ws}(r) = p \cdot (1 - p) \cdot p_1 \cdot p_2 \cdot p_3$$

Therefore, P_{ws} equals

$$P_{ws} = \int_0^1 2r P_{ws}(r) dr.$$

From the above analysis, we also know that the handshake between nodes x and y may fail due to the interference from the nodes in Area III. Thus, the failed period can range from $l_{rts} + 1$ to $l_{rts} + l_{cts} + l_{data} + l_{ack} + 4$. We also assume that it follows a truncated geometric distribution and use its mean length as T_{fail}. To take into account the greater effect of omni-directional transmissions of CTS packets that may well collide with ongoing transmissions, we use $l_{rts} + l_{cts} + 2$ as the lower bound of the distribution instead. Then the throughput can be calculated accordingly.

3.3.4   The *DTDR* Scheme

As to the *DTDR* scheme, referring to Fig. 3.2, it is clear that, only nodes in Areas I and II can interfere with the handshake between nodes x and y. However, in the *DTDR* scheme, nodes are more vulnerable to the transmissions from other nodes in these areas than they are in the *DTOR* scheme, because nodes receive omni-directionally only if they are in the *wait* state. To take the higher vulnerability into account, we use $l_{rts} + l_{cts} + 2$ as the lower bound for the distribution of T_f while the upper bound remains to be $l_{rts} + l_{cts} + l_{data} + l_{ack} + 4$.

On the other hand, because nodes x and y are immune to the transmissions from nodes in Areas III, IV, and V, and because concurrent transmissions can go on unobstructed

in these areas, we introduce a spatial reuse factor $\gamma(r)$ for this scheme in the calculation of $P_{ws}(r)$. The parameter $\gamma(r)$ is defined to be the number of possible concurrent transmissions in the combined region covered by nodes x and y, i.e.,

$$\gamma(r) = \min(\gamma_1(r), \gamma_2(r))$$

In the above expression, $\gamma_1(r)$ is the ratio between the total region covered nominally by nodes x and y and the actual region covered by the handshake between nodes x and y. If there is one handshake in Areas I and II, then in theory there can be possibly $\gamma_1(r)$ concurrent handshakes in the total area of Areas I to IV. Hence,

$$\gamma_1(r) = (S_I + S_{II} + S_{III} + S_{IV} + S_V)/(S_I + S_{II}).$$

On the other hand, $\gamma_2(r)$ can be explained as follows: There are on average $N \cdot (S_{III} + S_{IV} + S_V)$ nodes in Areas III to V and in theory they can form a maximum of $\gamma_2(r) = N \cdot (S_{III} + S_{IV} + S_V)/2$ pairs of concurrent handshakes. To be conservative, we take the smaller value of $\gamma_1(r)$ and $\gamma_2(r)$ to estimate the spatial reuse benefit enabled by the *DTDR* scheme.

The above is a very crude estimation of the gain in spatial reuse for the *DTDR* scheme, because the area unaffected by the handshake between nodes x and y cannot be used fully by their neighbors. Still, for this scheme, P_{ws} can be adjusted as follows:

$$P_{ws} = \int_0^1 2r\gamma(r)P_{ws}(r)dr.$$

The other quantities needed to derive the throughput are the same as those for the *DTOR* scheme unless specified earlier.

61

3.3.5 Analytical results

In this section, we investigate only the case in which data packets are much longer than control packets and compare the performance of the *OTOR*, *DTOR*, *MTOR* and *DTDR* schemes. This warrants an RTS/CTS based collision avoidance handshake before actual data packet transmissions.

We present the results of two typical configurations for these schemes; similar results can be readily obtained for other configurations. In these configurations, τ denotes the duration of one slot and RTS, CTS, and ACK packets last 5τ, while a data packet lasts 100τ. In configuration one, both omni-directional and directional transmissions have the same gain and thus we have $N' = N$. In configuration two, directional transmissions have higher gain than omni-directional transmissions and $\gamma = 1.5$. In this case, $N' = 1.5^2N = 2.25N$.

For each configuration, we derive the maximum achievable throughput when the antenna beamwidth changes from $\theta = 15°\left(\frac{\pi}{12}\right)$ to $\theta = 120°\left(2\pi/3\right)$ in increment of $\theta = 15°\left(\frac{\pi}{12}\right)$. The results for configurations one and two are shown in Figs. 3.5 and 3.6, respectively.

Figs. 3.5 and 3.6 clearly show that the *DTDR* scheme maintains the highest throughput among these schemes, even with the increase of transmission and receiving beamwidth. Two factors contribute to the superiority of the *DTDR* scheme. One is the significant increase in spatial reuse, because only a small area is covered by the transmissions between two nodes engaged in a handshake according to the analysis. The other is the much reduced interference from those nodes that are not aware of the handshake because of directional receiving.

Even though the *DTDR* scheme does not ensure perfect collision avoidance, the directional reception capability makes the receiving node immune to the transmissions from
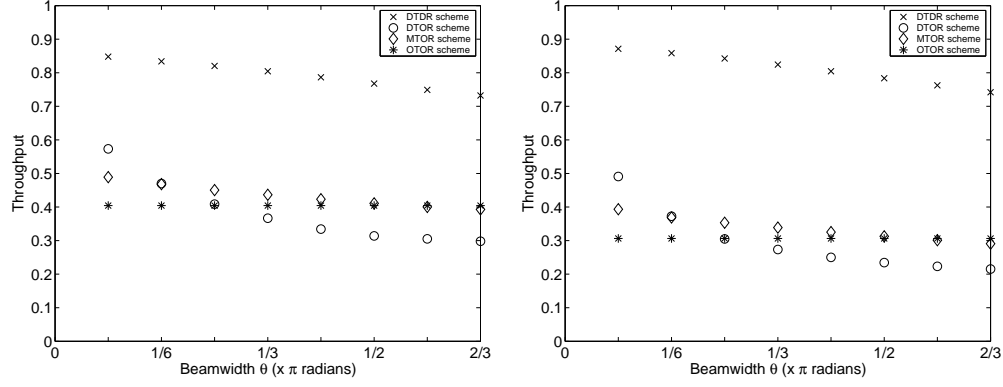
(a) Throughput (N=3)

(b) Throughput (N=5)
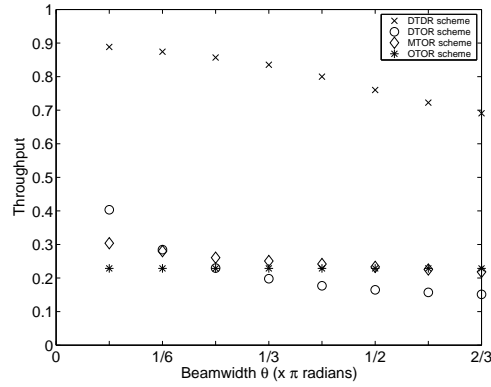
(c) Throughput (N=8)

Figure 3.5: Throughput comparison when omni-directional and directional transmissions have equal gain ($l_{rts} = l_{cts} = l_{ack} = 5\tau, l_{data} = 100\tau, \gamma = 1$).
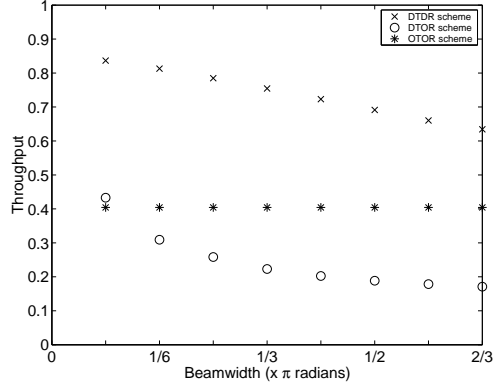
(a) Throughput (N=3)
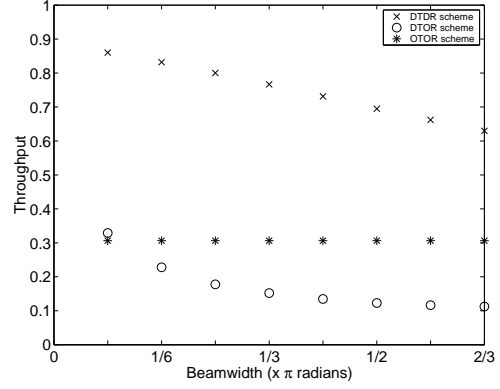
(b) Throughput (N=5)
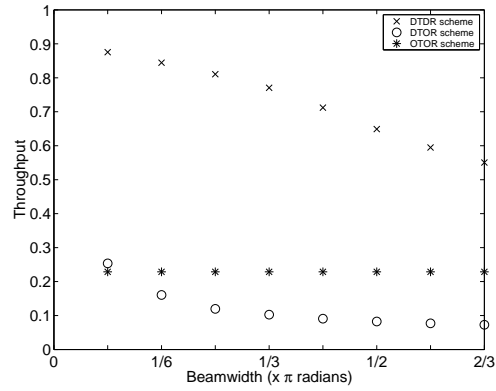
(c) Throughput (N=8)

Figure 3.6: Throughput comparison when the gain of directional transmissions equals one and a half times the gain of omni-directional transmissions ($l_{rts} = l_{cts} = l_{ack} = 5\tau, l_{data} = 100\tau, \gamma = 1.5$).

many other nodes in Areas III, IV, and V after it transmits a CTS packet. Hence, in terms of avoiding collisions, the *DTDR* scheme is as good as or even better than the *OTOR* scheme, which silences all the neighbors around both a sender and a receiver.

Another significant advantage of the *DTDR* scheme is that its performance does not degrade with the increase of competing nodes within a neighborhood when antenna beamwidth is narrow. Instead, it even has a slight increase in throughput. This can be explained as follows: When the number of nodes is small, spatial reuse may be not utilized to its full advantage because some nodes may have to stay idle when all of their neighbors are engaged. This is not due to collision avoidance, but due to the scarity of nodes. Hence, when more nodes are around, the effect of spatial reuse is more conspicuous and one-hop throughput increases accordingly. However, when antenna beamwidth increases, spatial reuse is reduced and throughput still degrades with the increase of N as people usually expect.

The results in Figs. 3.5 and 3.6 also show that, as expected, the performance of the *DTOR* and *DTDR* schemes degrades when directional transmissions have higher gain than omni-directional transmissions. This is a direct consequence of the fact that the higher gain of directional transmissions leads to more interference at nodes receiving in omni-directional mode. However, it is clear that the *DTDR* scheme is superior to the *OTOR* scheme in all cases, and the same conclusion can be derived from the results of simulation experiments described in Section 3.4. This helps to justify our approach of having each node consider as its neighbors those nodes that it hears through their omni-directional beacon transmissions.

3.4   Simulation Results

This section describes the results of computer simulations used to investigate the performance of the popular IEEE 802.11 DFWMAC protocol, which is labeled as *OTOR* in this section, and its variants corresponding to three directional collision avoidance schemes. The directional schemes considered are the *DTOR*, *MTDR* and *DTDR* schemes.

Most of the simulation settings remain the same as described in Section 2.4. We have implemented the directional schemes with the assumption that there is a neighbor protocol that can actively maintain a list of neighbors as well as their locations. In addition to evaluating the performance of the directional collision avoidance schemes with ideal directional antennas, we have also evaluated the performance of these schemes with directional antennas that generate side lobes in directional transmissions as shown in Fig. 3.1.

3.4.1   Performance evaluation with ideal directional antennas

Nodes are uniformly distributed in circles or rings and to avoid some extreme cases, we only use network topologies that satisfy the following conditions:

- For the inner N nodes, each node should have at least 2 neighbors and at most $2N - 2$ neighbors.

- For the intermediate outer $3N$ nodes, each node should have at least 1 neighbor and at most $2N - 1$ neighbors.

Fig. 3.7 illustrates a sample network topology used in our simulations when $N = 5$.
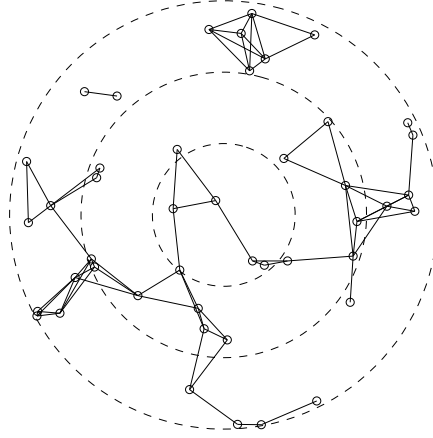
Figure 3.7: A sample network topology when $N = 5$

We run simulation programs with $N = 3, 5, 8$ with beamwidth $\theta = 30°, 90°, 150°$ respectively. We generate 50 random topologies that satisfy the uniform distribution and then get averaged throughput and delay for the N nodes in the innermost circle of radius R for each configuration.

The results for the case in which omni-directional and directional transmissions have equal gain are shown in Figs. 3.8 and 3.9. The results for the case in which directional transmissions have higher gain than omni-directional transmissions and $\gamma = 1.5$ are shown in Figs. 3.10 and 3.11.

In Figs. 3.8–3.11, the vertical lines show the range of throughput/delay achieved by each scheme, i.e., *mean ± standard variance*. The lines are shifted a bit for clarity.

At first, it can be seen that the throughput of the *DTOR* scheme does not degrade much in a relatively large range of transmission beamwidths when N is small and the throughput of the *MTOR* scheme degrades very little regardless of transmission beamwidth. This is a bit different from the prediction in the analysis where the throughput of both schemes de-
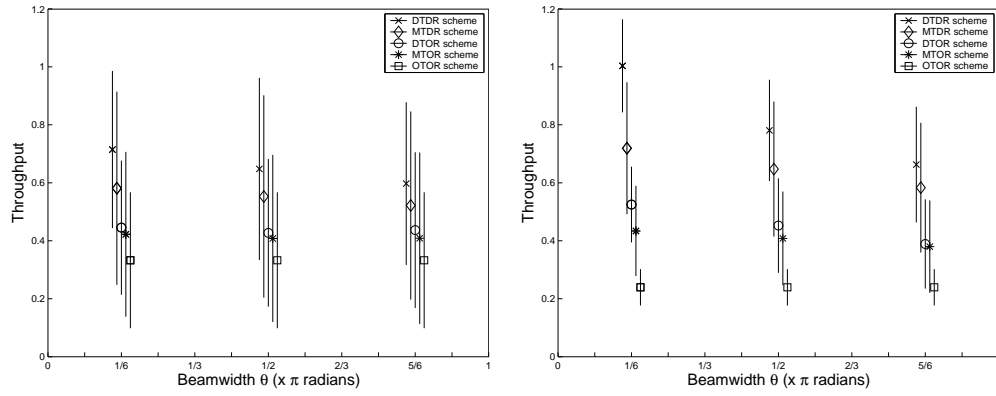
(a) N=3

(b) N=5
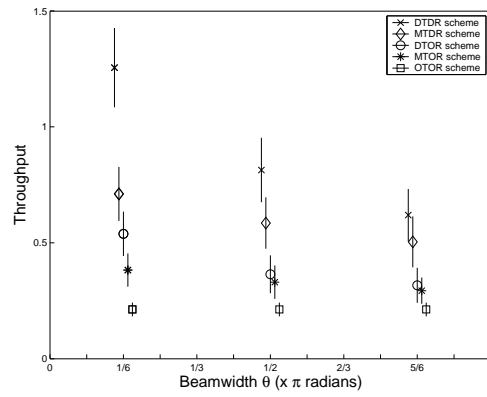
(c) N=8

Figure 3.8: Throughput comparison – equal gain
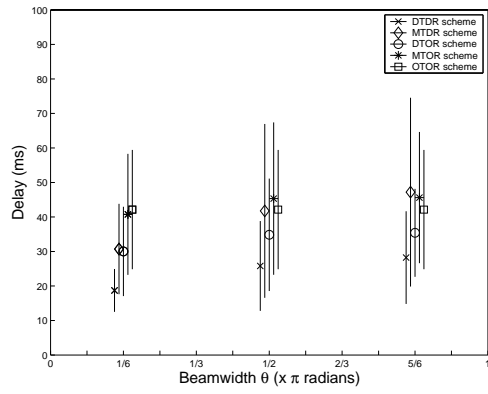
(a) N=3
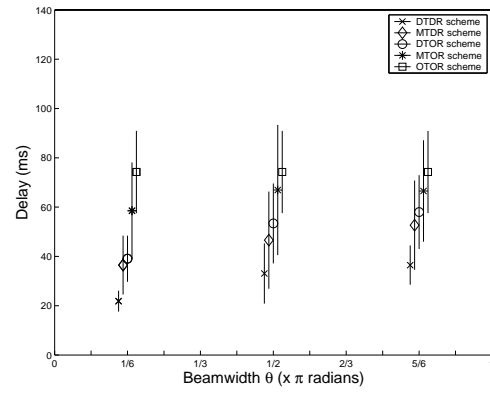
(b) N=5
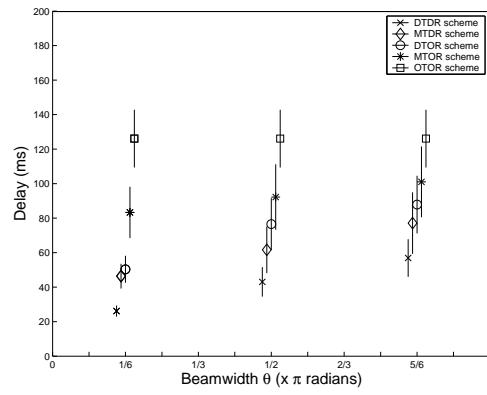
(c) N=8

Figure 3.9: Delay comparison – equal gain
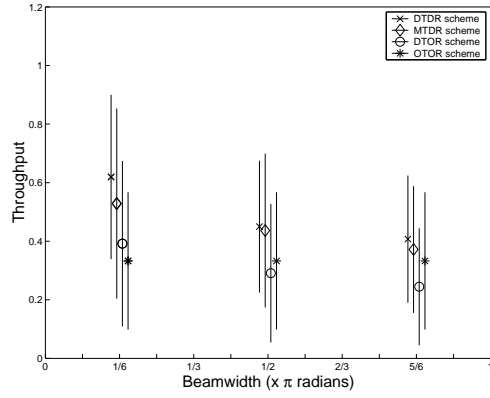
(a) N=3
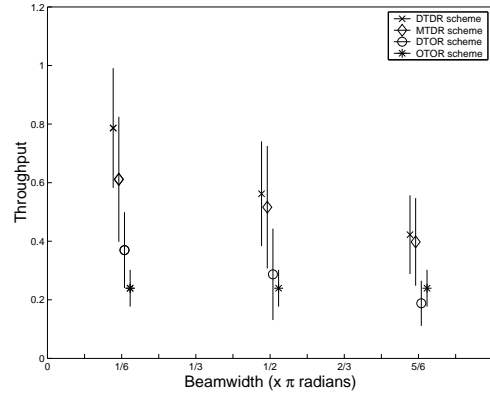
(b) N=5
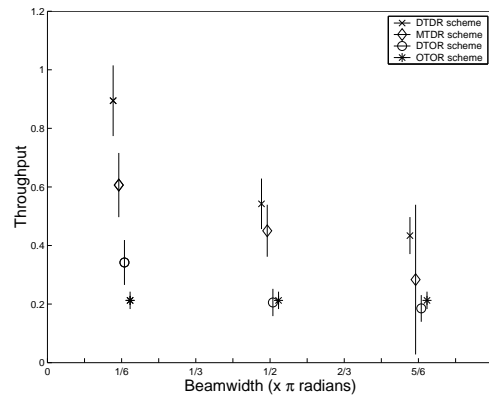
(c) N=8

Figure 3.10: Throughput comparison – higher gain ($\gamma = 1.5$)

70
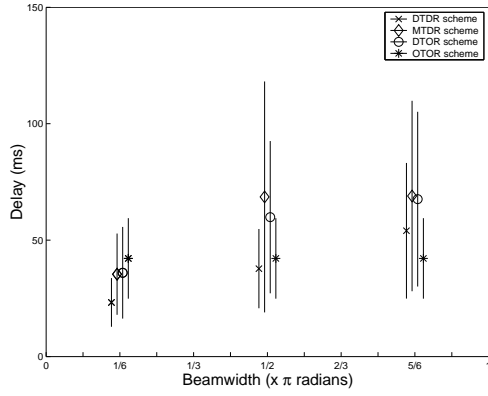
(a) N=3
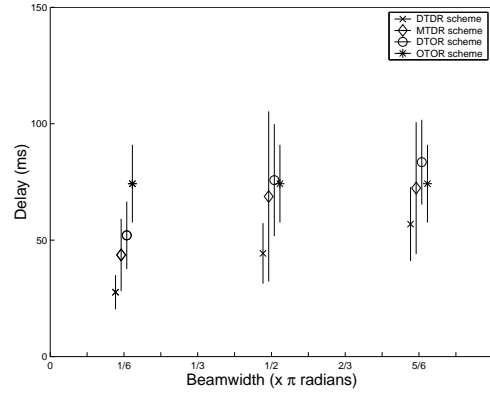
(b) N=5
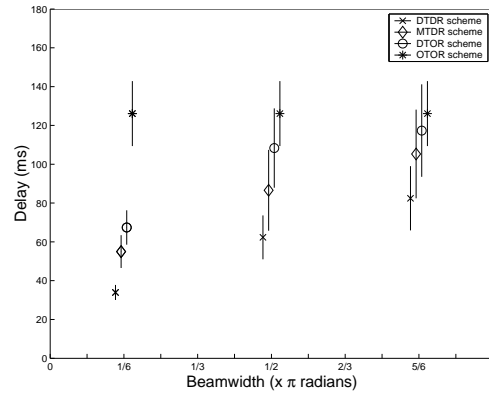
(c) N=8

Figure 3.11: Delay comparison – higher gain ($\gamma = 1.5$)

grades gradually with the increase of beamwidth. This can be explained as follows: The analysis does not take into account the fact that the number of nodes in a certain region can only be an integer. Instead, an infinite division of neighbors is implicitly assumed.

In the analytical model, it seems that the effects of a node vary in regions, which is clear from the derivation of the probability that one node succeeds in completing a four-way handshake with the other node. In the derivation, transmissions of the sending and receiving nodes are vulnerable to nodes from different regions for different periods. On the contrary, in simulations (and in reality) nodes have a definitive influence on the handshake of a pair of sending and receiving nodes. Referring to Fig. 3.2, for example, a node is either in Area III or V, but not in both and nor can it exert separate effects in these two areas. Thus, suppose that the neighbors of a node are distributed evenly around it, say three, it does not make any difference if this node transmits with beamwidth of $30°$ or $90°$.[2] Despite the inaccuracies of the analytical model, the simulation results still largely agree with what is predicted in the analytical model. That is, the *DTDR* scheme performs the best among all these schemes and its performance does not degrade even for large values of N as predicted in the analysis when antenna beamwidth is narrow. The results also show that the *MTDR* scheme outperforms the *DTOR* scheme, which indicates that the directional receiving capability can boost performance significantly.

Without directional receiving, a scheme with mixed transmissions (*MT* scheme) performs worse than a scheme with only directional transmissions (*DT* scheme). This is because omni-directionally transmitted CTS packets make almost all the nodes around the receiver

---

[2]In reality, it is more desirable to transmit with narrower beamwidth, because signal energy is more concentrated and a higher signal-to-noise ratio can be achieved, though physical layer impairment other than Gaussian white noise is not modeled in the simulations.

defer their access to the shared channel or interfere with the ongoing handshake around the nodes that transmit CTS packets. Such conservative collision avoidance can largely nullify the benefits of spatial reuse and an all-directional scheme such as *DT* is shown to perform much better than *MT* when both schemes use only directional transmission capability of antenna systems.

However, when directional receiving is used, even though CTS is transmitted omni-directionally, the handshakes of those nodes that have turned their receiving to other directions are not affected. Hence, the *MTDR* scheme can outperform the *DTOR* scheme in this case, although its performance is still inferior to the *DTDR* scheme because of the reduced spatial reuse.

It is also clear that, when beamwidth becomes wider, the performance of the *DTDR* scheme degrades faster when N becomes larger. This shows that when networks are dense, the performance of a directional scheme is more influenced by the transmission/reception beamwidth.

It should be noted again that, because correct collision avoidance is not enforced in the IEEE 802.11 MAC protocol, collisions of data packets can still occur and hence the *OTOR* scheme cannot achieve the same performance predicted in the analysis, which assumes correct collision avoidance. It is for this reason that the *DTOR* scheme performs better than the *OTOR* scheme, even when wider beamwidths are used.

When comparing the results shown in Figs. 3.8 and 3.9 with those in Figs. 3.10 and 3.11, it is clear that higher directional transmission gains can have negative effects on both throughput and delay. This is because a node's directional transmissions interfere with more

Table 3.1: Comparison of percentage of ACK timeouts in different schemes ($\gamma = 1$)

	DTDR			MTDR		
	30°	90°	150°	30°	90°	150°
N = 3	0.04±0.02	0.06±0.05	0.06±0.04	0.03±0.02	0.04±0.04	0.06±0.06
N = 5	0.05±0.01	0.08±0.03	0.11±0.05	0.04±0.01	0.08±0.04	0.13±0.07
N = 8	0.07±0.01	0.14±0.02	0.19±0.04	0.07±0.02	0.15±0.03	0.22±0.05

	DTOR			OTOR
	30°	90°	150°	N/A
N = 3	0.38±0.17	0.35±0.19	0.33±0.18	0.29±0.21
N = 5	0.46±0.07	0.45±0.11	0.44±0.12	0.39±0.10
N = 8	0.55±0.03	0.58±0.04	0.57±0.05	0.46±0.05

Table 3.2: Comparison of percentage of ACK timeouts in different schemes ($\gamma = 1.5$)

	DTDR			MTDR		
	30°	90°	150°	30°	90°	150°
N = 3	0.06±0.03	0.09±0.05	0.11±0.07	0.06±0.04	0.14±0.11	0.21±0.13
N = 5	0.08±0.02	0.13±0.05	0.17±0.06	0.10±0.04	0.20±0.09	0.30±0.10
N = 8	0.11±0.01	0.19±0.02	0.23±0.04	0.14±0.03	0.32±0.05	0.40±0.05

	DTOR			OTOR
	30°	90°	150°	N/A
N = 3	0.42±0.20	0.39±0.21	0.40±0.18	0.29±0.21
N = 5	0.51±0.11	0.51±0.15	0.51±0.15	0.39±0.10
N = 8	0.63±0.03	0.65±0.05	0.60±0.06	0.46±0.05

nodes, which translates into a reduction in spatial reuse, given that more nodes spend more time in the wait state after perceiving the channel busy.

We have also collected statistics about data packet collisions in the same way as we do in Section 2.4. The results are shown in Tables 3.1 and 3.2.

We find that the schemes with narrow receiving beamwidth have far smaller data packet collision ratios than the schemes without directional receiving. Hence, with directional

receiving, the adverse effects of hidden terminals are almost completely canceled, leading to much higher throughput. It can also be noted that higher directional transmission gain leads to higher data collision ratio due to the increased interference it introduces. Therefore, directional antenna systems that are able to transmit and receive with narrow beamwidth while having the capability to reduce the power of directional transmissions are much more desirable than other variants.

3.4.2 Impact of side lobes

As discussed in Section 3.3.1, we conjecture that side lobes in directional transmissions should not have much effect on throughput if the side lobe level is low enough and carrier sensing threshold is raised. Hence, we have implemented the directional antenna model shown in Fig. 3.1 in GloMoSim and conducted some simulations. In our simulations, we also set the interference range of side lobes $d = R/\gamma$. Other configurations remain the same as the case when directional transmissions have higher gain, i.e., $R' = \gamma R$. Simulation results are shown in Figs. 3.12 and 3.13 and we can see clearly that the presence of side lobes only causes larger variance for access delay when N is small, and its effect is negligible in all the other cases.

3.5 Conclusion

In this chapter, we have presented the first analytical model for directional collision avoidance protocols which takes into account both directional transmission and reception capabilities and the possibility of having different gains in omni-directional and directional
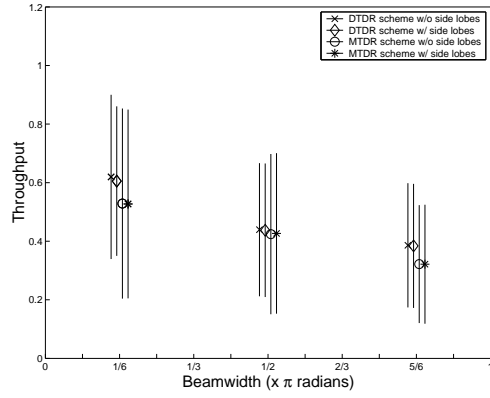
(a) N=3
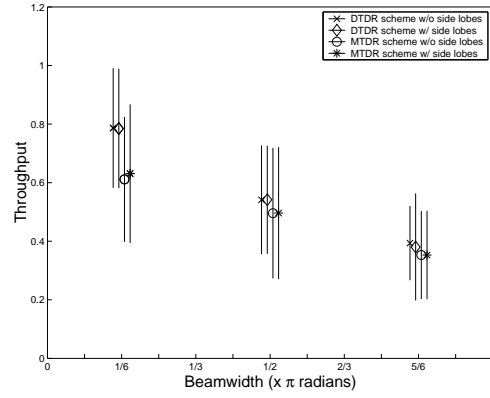
(b) N=5
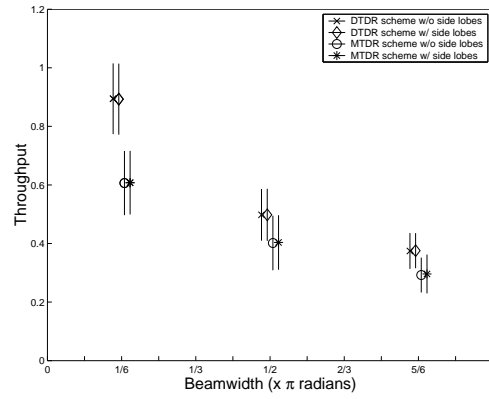
(c) N=8

Figure 3.12: Throughput comparison – higher gain ($\gamma = 1.5$) with side lobes
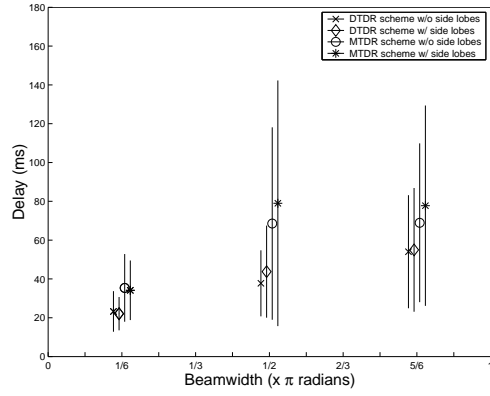
(a) N=3
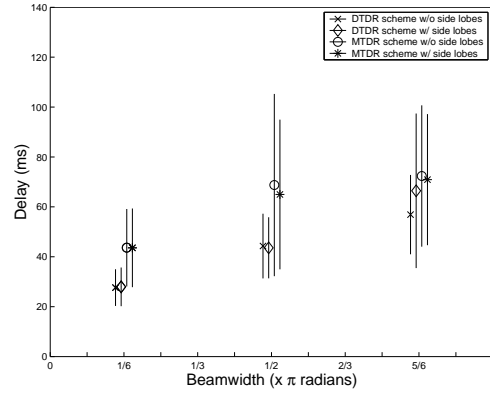
(b) N=5
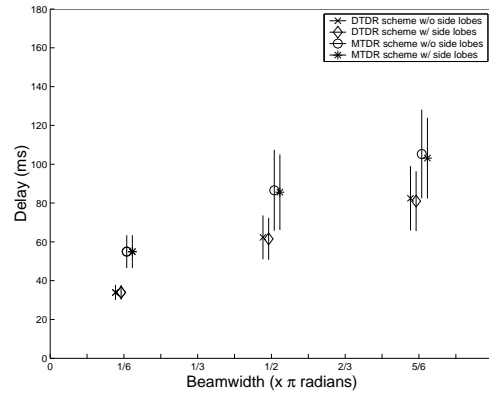
(c) N=8

Figure 3.13: Delay comparison – higher gain ($\gamma = 1.5$) with side lobes

77

transmissions. The analytical results show that the scheme in which all transmitting and receiving are done directionally can achieve much higher throughput than any other scheme that combines directional and omni-directional transmissions or receptions. The all-directional scheme maintains high spatial reuse and largely cancels the interference from hidden terminals due to imperfect collision avoidance. Furthermore, the one-hop throughput of the all-directional scheme does not degrade with the increase of competing nodes within a region, which shows that the all-directional scheme is also much more scalable in dense ad hoc networks. It is also shown that higher directional transmission gain can have negative effects on the performance of directional collision avoidance schemes due to the increased interference range and reduced spatial reuse.

Extensive simulations of the popular IEEE 802.11 MAC protocol and its directional variants validate the analytical results. The very low ratio of data packet collisions in the schemes with directional receiving confirms that directional receiving can cancel out almost all the adverse effects of hidden terminals which seem to be the throughput bottleneck even for conservative collision avoidance scheme as exemplified in the IEEE 802.11 MAC protocol.

Simulation results also show that side lobes have little effect on throughput if side lobes are reasonably suppressed and carrier sensing threshold is raised to make nodes less sensitive to channel activities. Augmented with the work to be presented in the next chapter that shows broadcast traffic does not degrade much the performance of directional collision avoidance schemes, it is argued that an all-directional scheme is very attractive and practical for ad hoc networks. It attains much better throughput and delay than the other schemes, and the neighbor protocol that it needs to obtain location information of neighboring nodes can be

implemented using very simple methods, without degrading its performance significantly.

In practice, some form of power control to achieve similar gains for both omni-directional and directional transmissions is desirable to take full advantage of the antenna systems. It is also possible to use power control to reduce both interference and energy consumption. Interesting areas of future research include analyzing the impact of eliminating omni-directional transmissions and receptions altogether by means of a directional beaconing mechanism, and comparing the performance of such a scheme against schemes that rely on omni-directional beaconing.

Chapter 4

Directional MAC Schemes with Broadcast Traffic

4.1 Introduction

Broadcasting is used extensively in routing protocols for multi-hop ad hoc networks. For example, route discovery (including route queries and replies) and neighbor information exchange rely on broadcasting which is much more cost effective than sending copies of unicast data packets to each interested node, especially when high reliability is not required. Because simple flooding wastes considerable channel resources and can lead to excessive collisions, how to disseminate information from one source node to all the other nodes in the network with a small number of broadcast packets has become an intensely researched topic in recent years. Some work (e.g. [74–76]) focus on efficient time-slot assignments in a channel access environment similar to dynamic-TDMA, which is not readily applicable to contention-

based random access environment with no time-slotted structure. Some recent work addresses the problem of efficient broadcast in ad hoc networks without global time synchronization.[1] For example, Ni et al. [77] propose and evaluate several approaches that make use of distance information, location information, or clustering to reduce the number of copies of broadcast packets sent. Lim and Kim [78] prove that the problem of finding the minimum-cost flooding tree is similar to the minimum connected dominating set (MCDS) and that it is an NP-complete problem. Then the authors propose self-pruning and dominant pruning schemes that utilize topology information to approximate the theoretical minimum of broadcast packets sent from one source to all the other nodes. In fact, there is already considerable work on this topic, and Williams and Camp [79] have given a detailed comparison of these broadcasting techniques for multi-hop ad hoc networks.

On the other hand, the goals for reliable broadcasting are different from that of best-effort broadcasting which is to reduce the redundancy of broadcast packets. In contention-based MAC protocols such as IEEE 802.11 MAC protocol [16], the usual RTS/CTS based collision avoidance handshake does not work for sending broadcast packets as it is one-to-many communication. Hence, broadcast packets are sent whenever a node senses the channel idle, which may collide with the transmissions from hidden nodes [27]. Tang and Gerla [80, 81] address the unreliable MAC layer broadcast problem by extending the RTS/CTS collision avoidance scheme. Tang and Gerla [80] first propose broadcast support multiple access (BSMA) protocol, which depends on a node's direct sequence (DS) capture capability to receive the CTS with the strongest signal without being affected by other CTS packets. If the neighbors

---

[1]Global time synchronization is necessary and sufficient for a time-slotted system, which is not easily achievable in multi-hop ad hoc networks.

that have sent CTS packets fail to receive the ensuing broadcast data packets, they will send negative acknowledgments (NACKs) to notify the source node. Though it is not guaranteed that all the neighbors of a source node can receive the broadcast packet from it, this scheme indeed improves the broadcast packet delivery ratio thanks to the RTS/CTS collision-avoidance procedure. The broadcast medium window (BMW) proposed by the the same authors [81] provides more reliable broadcast support, although with much more complexity. In this scheme, the MAC protocol needs to maintain a neighbor list, a list of transmitted packets and a list of sequence numbers of received packets. A node sends its packets to its neighbors in a round robin flavor with the usual carrier sense multiple access with collision avoidance (CSMA/CA) scheme, i.e., an RTS-CTS-data-ACK handshake. The major difference lies in the fact that sequence numbers are exchanged in the RTS/CTS handshake. In this way, a node can send the next packet if the intended receiver happens to have overheard and received the previous packets sent by this node.

However, all these enhancements to broadcasting in multi-hop ad hoc networks have been analyzed when there exists only broadcast traffic and some investigations are limited to low and medium traffic load so as to isolate the effects of loss due to contentions. In an operational ad hoc network, a mix of unicast and broadcast traffic is quite common and it is important to investigate both the effects of broadcast traffic on unicast traffic and the performance of broadcast traffic in the presence of unicast traffic of these proposed schemes. This motivates our work in investigating the interaction between unicast and broadcast traffic.

It would be just a simple extension to the work done by Williams and Camp [79] if we just investigated these schemes with a mix of unicast and broadcast traffic. Instead, we

are interested in exploring another dimension in the solution space for broadcast in ad hoc networks: *directional antennas*. As prior chapters have shown, directional MAC protocols can increase the performance of unicast traffic. However, no prior work has investigated how the use of directional antennas may influence the performance of both unicast and broadcast traffic which is the focus of this chapter.

In Section 4.2, we present the results of simulation experiments with the *DTOR*, *MTOR* and *OTOR* schemes discussed in Chapter 3. we show that throughput and delay can vary widely even in networks in which nodes are uniformly distributed. Thus, it is very important to conduct performance evaluation with different network configurations and to consider both mean value and variance of interested metrics such as throughput and delay. We also show that the MAC scheme that maximizes the use of directional transmissions (the *DTOR* scheme) has the best throughput-delay performance among all the schemes for both unicast and broadcast traffic. Finally, in Section 4.3, we summarize the findings in this chapter and propose some future work.

4.2   Simulation Results

In our simulation experiments, we investigate both the effects of broadcast traffic on unicast traffic and the performance of broadcast traffic in the presence of unicast traffic. The latter is different from the work reported in some literature where performance of broadcast traffic is investigated isolatedly without the presence of unicast traffic. We use the same simulation models as those in Section 3.4 and we show the results for two sets of simulations.

4.2.1 Unicast traffic performance in the presence of broadcast traffic

In our simulation experiments, each node in the innermost circular region of radius R is a constant bit rate (CBR) generator that continuously generates unicast data packets and broadcast data packets alternately. For unicast packets, the destination node is chosen randomly from the node's neighbors. The size of a unicast data packet is 1460-byte and the size of a broadcast data packet is 500-byte, about one third the size of a unicast data packet. We use r to denote the ratio of the number of broadcast data packets generated to the total number of data packets generated.

We vary the number (N) of nodes in the innermost circular region of radius R as well as r to obtain three metrics. The first one is unicast throughput which is the aggregate throughput contributed by the innermost N nodes sending unicast data packets that are acknowledged. The second one is broadcast throughput. We count the number of broadcast packets from the innermost N nodes that are received by the neighbors of the innermost N nodes as well as these nodes themselves.[2] Then we divide it by the number of broadcast data packets that can be transmitted in the total simulation time to obtain the normalized broadcast throughput. The third one is the average delay of unicast data packets. In addition, for the two schemes (*DTOR* and *MTOR*) that utilize directional transmissions, we also investigate their performance under different values of beamwidth θ: 30°, 90° and 150°.

For each configuration, we generate 30 random topologies and for each topology we run the simulations with three different seed numbers. For each run, we calculate the interested metrics and take the average over the three different seed numbers. Then we calculate the

---

[2] A broadcast packet may be counted more than once if more than one node receive it correctly.

84

mean values and standard variance of these metrics. The results are shown in Figures 4.1–4.4. In these figures, the vertical lines show the range of metrics achieved by each scheme, that is, *mean ± standard variance*. These lines are shifted a bit for clarity.

Figures 4.1–4.2 show that all the throughput of the three schemes degrades rather elegantly even when on average each node sends 30% of broadcast traffic. Here the only exception is that the throughput of *MTOR* scheme increases a bit for $N = 5, 8$ when r increases from 0 to 0.1. It seems that in these cases a small percentage of broadcast traffic helps to interrupt some nodes' long waiting time for collision avoidance and thus nodes are more aggressive in access to the shared channel. However, for other cases, broadcast traffic can degrade unicast throughput almost definitely. In addition, for small values of N (such as 3), the three schemes perform almost the same, considering both mean and standard variance. When N increases, the *DTOR* scheme with small beamwidth θ performs indisputably much better than the other two schemes. This can be explained as follows. At first for unicast traffic, when the network becomes more congested, it is very difficult for a pair of sending and receiving nodes to get coordinated with their one-hop and two-hop neighbors in the *OTOR* scheme. If coordination is not achieved, then their handshake may be very probably disrupted by the omni-directional transmissions of neighboring nodes. Even if coordination is achieved, all their one-hop and two-hop neighbors are prohibited from transmitting and spatial reuse is greatly reduced. The same reason applies to *MTOR* scheme due to the omni-directional transmission of CTS packets. In the *DTOR* scheme, transmissions are confined to much smaller regions and multiple flows may coexist at the same time. When the network is less congested, tradeoff between collision avoidance and spatial reuse is much more balanced and all schemes work similarly.

For broadcast traffic in congested networks (large values of N), although all the three schemes use the basic CSMA to send broadcast packets, however as more unicast packets can be sent in the *DTOR* scheme, hence are broadcast packets. Accordingly, the *DTOR* scheme can achieve the highest broadcast throughput for large values of N.

It should also be noted that, the performance metrics can vary a lot even when the same uniform distribution is used throughout the simulation experiments, especially when N is small. Hence we argue that it is very important to experiment with enough network topologies before conclusions can be drawn, otherwise misleading results may be obtained.

Figures 4.3–4.4 show that the broadcast throughput increases almost linearly with the increase of percentage of broadcast traffic and with the number of innermost nodes. Besides, the *DTOR* scheme is shown to have larger marginal gain in broadcast throughput. This is due to the more aggressive channel access scheme in *DTOR* where the gain in spatial reuse outweighs the gain in conservative collision avoidance. It is also evident that the *DTOR* scheme has more distinct advantage when beamwidth is narrow, say $\theta = 30°$. We also find that the *DTOR* scheme with small beamwidth also achieves the least delay for unicast traffic as nodes spend less time in collision avoidance.

4.2.2 Broadcast traffic performance in the presence of unicast traffic

We still use the circle and ring topology with $R = 3$ and focus on the broadcast packets received as well as their delay recorded by the innermost N nodes under light to medium traffic load. We show the results when $r = 0.3$, i.e., about 30% of the packets sent by any node are broadcast packets, and $\theta = 30°$ when applicable. Table 4.1 shows the
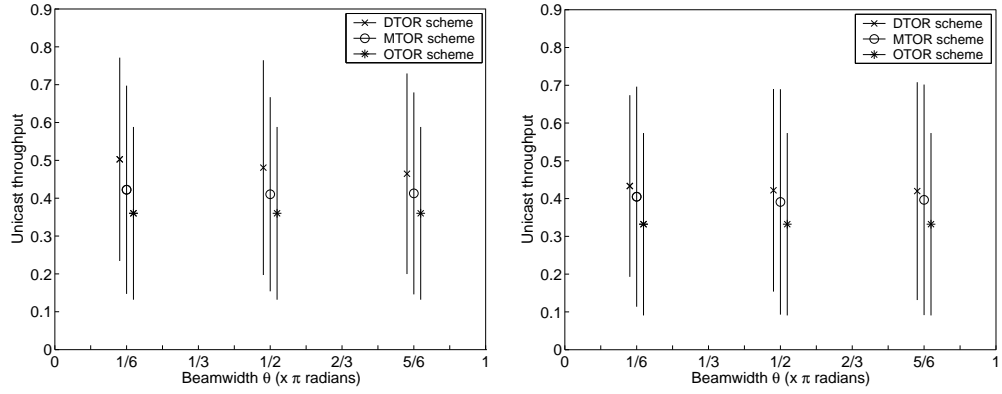
(a) $N = 3, r = 0.0$

(b) $N = 3, r = 0.1$
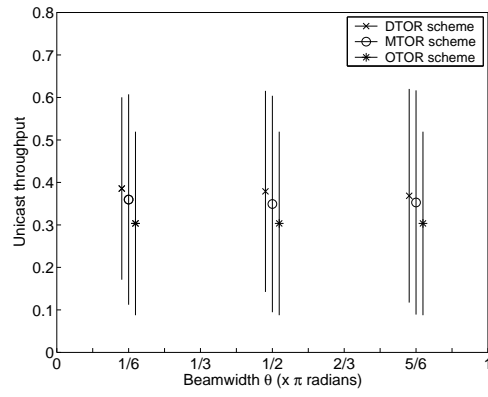
(c) $N = 3, r = 0.3$

Figure 4.1: Unicast throughput ($N = 3$)
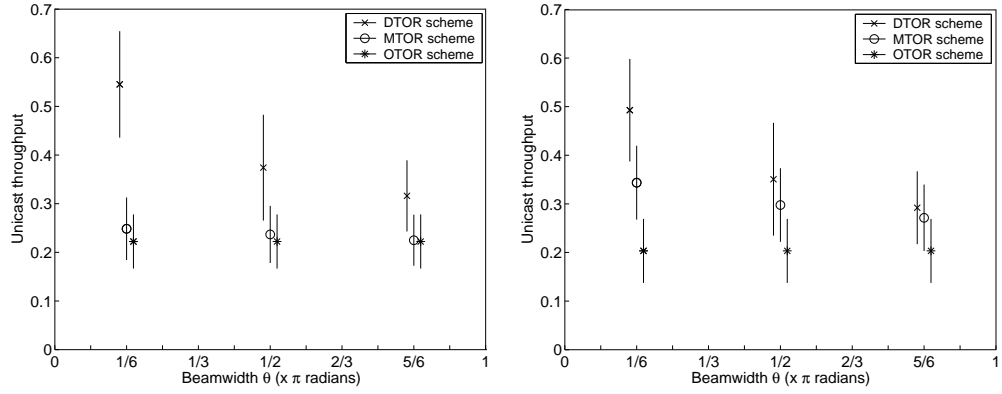
(a) $N = 8, r = 0.0$

(b) $N = 8, r = 0.1$
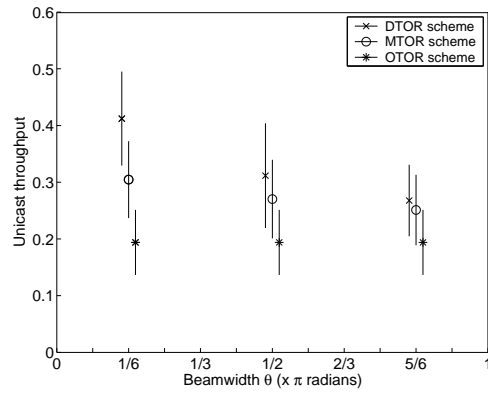
(c) $N = 8, r = 0.3$

Figure 4.2: Unicast throughput ($N = 8$)
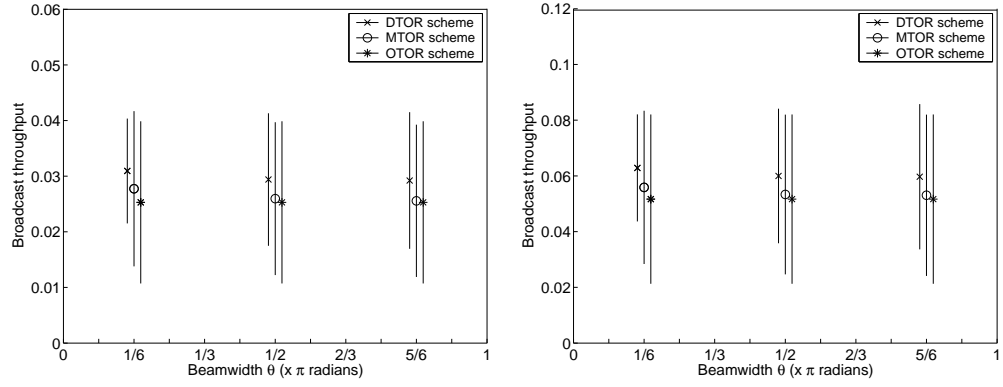
(a) $N = 3, r = 0.1$

(b) $N = 3, r = 0.2$
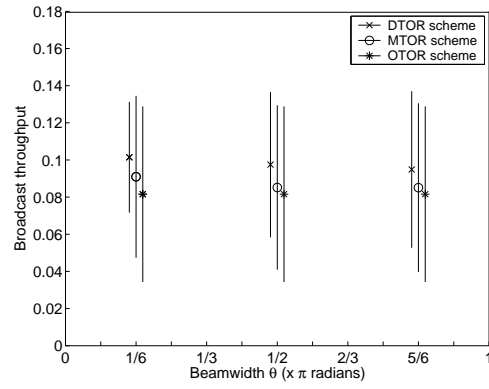
(c) $N = 3, r = 0.3$

Figure 4.3: Broadcast throughput ($N = 3$)
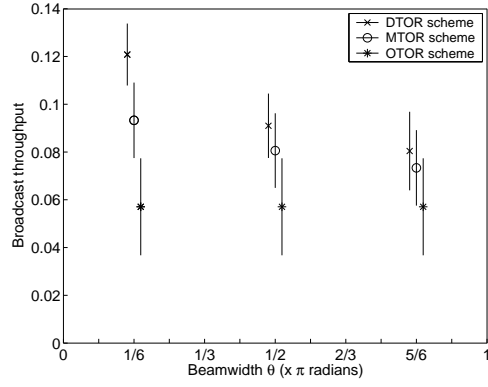
(a) $N = 8, r = 0.1$
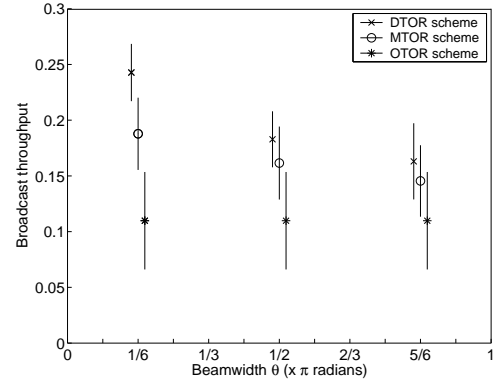
(b) $N = 8, r = 0.2$
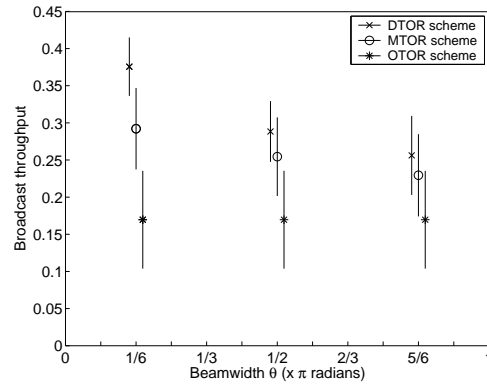
(c) $N = 8, r = 0.3$

Figure 4.4: Broadcast throughput ($N = 8$)

three configurations for different values of N in which the packet interarrival time is varied to change the offered load to the shared channel. Configurations one and two roughly correspond to light load and configuration three medium load. Table 4.2 shows the average delay and standard variance of broadcast packets observed by the innermost N nodes. Table 4.3 shows the average and standard variance of the delivery ratio of broadcast packets observed by the innermost N nodes. Here the ratio is the total number of broadcast packets received to the total number of broadcast packets generated by the innermost N nodes.

It can be observed that, with regard to delay in light load scenarios, in general *DTOR* has no distinct advantage over *MTOR* scheme, although both schemes perform better than the *OTOR* scheme; in medium load scenarios, in general *DTOR* has distinct advantage over the other two schemes in that both the average delay and variance are much smaller. The difference lies in the fact that unicast packets can be delivered much faster in the *DTOR* scheme whose channel access is more aggressive, while the other two use more conservative channel access scheme. We also note that the variance of all three schemes is quite large. Thus it again supports our argument that, in the simulation of ad hoc networks, it is very important to run simulations with different topologies extensively even when the topology distribution seems quite regular, such as the uniform distribution, otherwise it is very probable to get misleading results.

Table 4.3 also shows that in general the nodes can receive more broadcast packets in the *DTOR* scheme than the other two schemes. In summary, the *DTOR* scheme can achieve the best delay-throughput curve among the three schemes investigated.

Table 4.1: Packet interarrival time configurations

	conf 1	conf 2	conf 3
N=3	108ms	72ms	54ms
N=5	180ms	120ms	90ms
N=8	288ms	192ms	144ms

Table 4.2: Broadcast packet delay comparison (unit: ms)

	DTOR			MTOR			OTOR		
	conf 1	conf 2	conf 3	conf 1	conf 2	conf 3	conf 1	conf 2	conf 3
N=3, mean	11.7	24.6	113.5	6.8	18.1	251.3	11.5	80.7	480.1
N=3, std	4.0	11.0	185.9	1.8	27.5	560.5	5.4	236.2	672.5
N=5, mean	8.1	16.9	35.7	8.5	12.3	112.8	14.1	52.6	590.3
N=5, std	2.0	6.0	26.7	1.5	6.7	336.0	3.9	67.0	835.4
N=8, mean	7.5	9.3	17.3	10.3	11.4	26.2	16.3	39.8	569.6
N=8, std	1.2	1.5	3.7	1.8	2.9	22.3	4.1	25.1	778.2

Table 4.3: Broadcast packet delivery ratio comparison

	DTOR			MTOR			OTOR		
	conf 1	conf 2	conf 3	conf 1	conf 2	conf 3	conf 1	conf 2	conf 3
N=3, mean	0.32	0.95	2.27	0.23	0.82	2.08	0.19	0.72	1.75
N=3, std	0.15	0.48	1.13	0.17	0.58	1.25	0.15	0.51	0.98
N=5, mean	0.14	0.50	1.29	0.13	0.39	1.44	0.06	0.31	1.03
N=5, std	0.05	0.19	0.51	0.06	0.18	0.78	0.03	0.29	0.39
N=8, mean	0.10	0.31	0.67	0.08	0.25	0.87	0.03	0.13	0.66
N=8, std	0.03	0.07	0.13	0.04	0.09	0.52	0.01	0.12	0.27

4.3 Conclusion

In this chapter, we have investigated the effects of directional antenna on the performance of both broadcast and unicast traffic through extensive simulation experiments. We show that it is very important to experiment with different configurations, even when relatively regular network topologies are used and both mean value and variance of the performance metrics should be considered. We also show that the collision avoidance scheme that maximizes spatial reuse by making transmissions all directional achieves the best performance among the three schemes investigated. Its more aggressive channel access scheme helps to achieve higher throughput and reduced average delay for both unicast and broadcast traffic.

Though the proposed enhancements to broadcasting in the literature have not been incorporated in our investigation, it can be reasoned that the use of directional antennas as another dimension in the solution space to enhance broadcasting in multi-hop ad hoc networks, is orthogonal to the proposed enhancements reported in the literature [77–81] and using them in combination with all-directional collision avoidance may help to improve the performance even further. This is an interesting topic for future work.

Chapter 5

Fairness in MAC Protocols

5.1 Introduction

In the usual collision avoidance schemes, nodes have to back off random amounts of time and then attempt their channel access again when collisions do occur. Due to its stability and long-term fairness, the binary exponential backoff (BEB) is favored in most MAC protocols and notably is adopted in the IEEE 802.11 MAC protocol. The BEB scheme is very effective in collision resolution when nodes face the same or similar level of contention. This is because when collisions occur, nodes double their contention window[1] and back off accordingly which reduces the contention significantly.

However, there are short-term and medium-term fairness problems associated with this scheme, because a node resets its contention window to the minimum size when it succeeds in sending a data packet. Accordingly, the node that last succeeds is much more ag-

---

[1]Backoff timer is chosen from a uniform distribution that is bounded between 0 and the current contention window.

gressive by comparison in its next access to the channel and may monopolize the channel for a long time while other nodes suffer starvation. In fact, the fairness problem even exists in single-hop networks. For example, the problem is known as the *capture effect* in Ethernet. Some modifications to the BEB have been proposed in the past to address the capture effect, including the capture avoidance binary exponential backoff (CABEB) [82] and the binary logarithmic arbitration method [83]. However, these modifications cannot be applied directly to multi-hop wireless networks, because they assume that each node has the knowledge of other nodes' successes and collisions.

Bharghavan et al. [14] are among the first people to address the fairness problem in multi-hop wireless networks. Some schemes have been proposed to alleviate the fairness problem since then. These schemes can be roughly divided into two categories. In the first category the goal is to achieve max-min fairness [42,50,84]. To be specific, these schemes try to reduce the ratio between maximum throughput and minimum throughput of flows, at either node's level or link's level. In the second category, the approach used in fair queueing [85–88] for wireline networks is adapted to multi-hop ad hoc networks, taking into account the salient characteristics of such networks such as location-dependent contention, distributed coordination and possible spatial reuse [44–48]. In these schemes, the contention among nodes is abstracted into a flow contention graph. Fig. 5.1 shows an example of how this is done. Any two flows with adjacent vertices in the flow contention graph should not be scheduled to transmit at the same time. Each node decides its backoff time from the service tag (or start tag, depending on which fair queueing discipline for wireline networks is used for approximation) of their own flows and other flows in their local neighborhood or the whole

Node graph

Flow contention graph
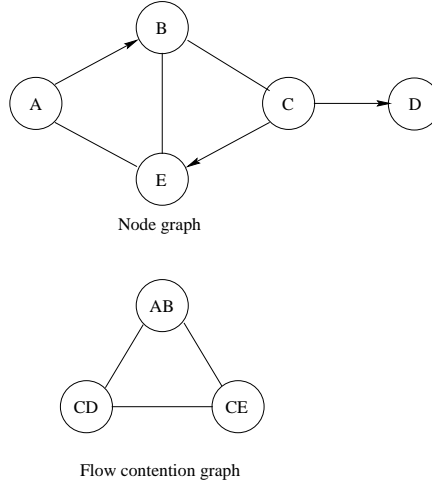
Figure 5.1: A simple network: node graph and flow contention graph

network. Usually, the flow with the earliest service tag (or start tag) needs to back off for the

minimum time so that its head-of-line (HOL) packet may be transmitted first.

The approaches in the second category are appealing because some service assur-

ance may be provided for ad hoc networks if they can approximate the fair queueing algo-

rithms used in wireline networks. In this chapter, we address two problems associated with

these schemes. The first is the coordination problem that results from the multi-hop charac-

teristic of ad hoc networks. The difficulty of multi-hop coordination can make these backoff-

based distributed fair queueing schemes less effective, which we discuss in Section 5.2. The

second problem is that the flow contention graph is not sufficient to model the contention

among nodes. By investigating how two competing flows share the available channel band-

width with different underlying network topologies but with the same contention graph, we

show that various degrees of fairness problems can take place. Section 5.3 presents the simu-

lation settings for two competing flows, in which we show that more than 10 different network

topologies can have the same flow contention graph even though there are only two flows in

these networks. Section 5.4 presents the results for TCP traffic, UDP traffic, as well as a mix of TCP and UDP traffic in these networks. It is shown that the fairness problem does exist in the IEEE 802.11 MAC protocol and TCP-based flows suffer more, because the acknowledgment traffic from TCP can have negative effects on both fairness and throughput. Section 5.5 summarizes our findings and motivates our work to be presented in Chapters 6 and 7.

5.2 Multi-Hop Coordination

Most existing distributed fair-queueing algorithms for ad hoc networks proposed in the literature modify the backoff scheme such that the sender of a flow with the minimum service tag (or start tag) within a contention region can be almost the first to transmit an RTS by backing off for the minimum amount of time. Here the contention region includes all the flows that may collide with the interested flow. However, due to hidden terminals, it need not be the case that the node with the minimum backoff time in a multi-hop network can transmit an RTS packet that succeeds. This can happen when the difference between the nodes' backoff times (measured in time slots for ease of discussion) is not large enough. In that case, another sender of a flow with the second-to-minimum service tag may also transmit an RTS, which leads to collisions.[2] This can be illustrated by the network depicted in Fig. 5.1, where we consider only two flows, AB and CE. Suppose that node A has the minimum service tag and node C has the second-to-minimum service tag, such that the difference in backoff slots of A and C is b. Suppose that both RTS and CTS last n slots and ignore the propagation delay. If b is between 1 and $n - 1$ (which is in fact $rts - 1$), then C's RTS will collide with A's RTS at

---

[2]The node transmits because each node maintains only flow contention information and does not necessarily know the underlying network topology.

B. In this case, the difference between the two flows is not large enough for the one with the earlier service tag to access the channel successfully. Even worse, a node with the second-to-minimum service tag may even win its access to the channel over a node with the minimum service tag. For example, consider the two flows AB and CD shown in Fig. 5.1. If flow AB has the minimum service tag for its HOL packet and the backoff difference between the two flows is less than $n - 1$, then C can send an RTS packet before CTS from node B arrives at C, which makes D reply with a CTS and B backoff, which means that C can transmit its HOL packet successfully to D, even though it does not have the minimum service tag. Hence, although all nodes implement the protocol faithfully, they fail to achieve the desired goal. This is called the *priority reversal* problem by Yang and Vaidya [89] though it is discussed in a different context.

Another problem is that sometimes the flow with the second-to-minimum service tag may be penalized in some cases due to the flow with the minimum service tag. For example, consider flows AB and DC in Fig. 5.1. Suppose that flow AB has the minimum service tag and DC has the second-to-minimum service tag. If the difference b is larger than $rts + 1$ and less than $rts + cts + data + ack + 1$, which is a large number, it is impossible for node D to initiate a successful handshake before nodes A and B finish theirs. Hence the transmission from any node like D is almost doomed to failure, even though it has backed off for such a long time. In addition, if the flow AB cannot finish sending its HOL packet in due time and the difference b between flow AB and flow DC drops below $rts + cts - 1$, then repeating collisions may occur if precautions are not taken.

It is evident from the above discussion that the required coordination among multi-
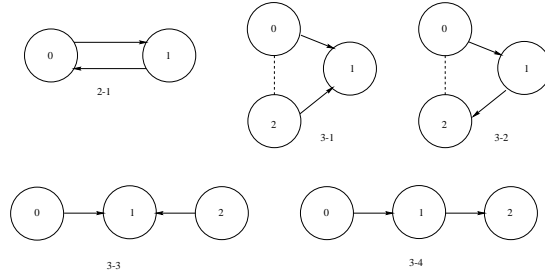
Figure 5.2: Networks with 2 or 3 nodes

hop nodes makes a backoff scheme derived from rankings of service tags less effective than what is expected, especially when only a flow contention graph is maintained and used in each node.

5.3 Two Competing Flows

We are interested in how the channel bandwidth is divided among two competing flows under different spatial contention characteristics and traffic patterns (either TCP or UDP). We do not consider two competing flows that originate from the same node, because a node can perfectly avoid sending packets of the two flows at the same time. Furthermore, this is a local (or inner-node) scheduling problem that can be handled by the node itself. All the possible configurations are shown in Figs. 5.2 and 5.3, in which a dashed line means that two nodes can hear each other's transmissions and an arrow indicates an active flow between two nodes. Nodes without any line in-between are hidden from each other. It is surprising to observe that there are so many variations even for such a simple case. We need to investigate how the different underlying network topologies can lead to various degrees of fairness problem via simulations which are presented in next section.
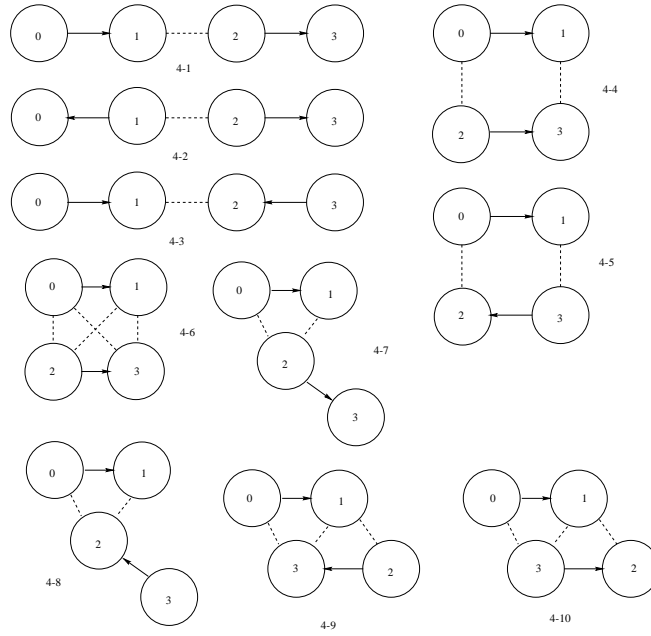
Figure 5.3: Networks with 4 nodes

It should also be noted that, in this and subsequent chapters, we do not consider the case of multiple competing flows. Due to the nature of the shared channel, at most one flow can succeed at any time and the presence of more competing flows just causes more collisions and backoffs, and makes one flow more difficult to succeed. This can have two different effects. On the one hand, the fairness problem may be mitigated, because a flow cannot consistently beat all the other competing flows for a long time, so the channel may be shared more evenly among competing flows. On the other hand, the fairness problem may be aggravated. A few flows with favorable locations can take turns to grab the channel while the other flows may be totally denied access to the channel. Hence, although no node monopolizes the channel in the presence of many flows, the fairness problem still persists. Whether multiple flows mitigate or aggravate the fairness problem largely depends on the network topologies and distribution of active flows. Given that the fairness problem has not been solved for even two flows, and

that a detailed analysis of fairness under several flows is very involved, this thesis focuses on the case of two competing flows only.

5.4 Simulation Results and Discussions

We have done simulations with the network configurations shown in Figs. 5.2 and 5.3 using GloMoSim 2.0 [73]. The raw channel bandwidth is 2Mbps and the underlying MAC protocol is IEEE 802.11 with direct sequence spread spectrum (DSSS) physical layer parameters. Table 2.1 shows the detailed parameters used throughout the simulations.

The simulations with the existing IEEE 802.11 MAC protocol include three sets. The first set shows the results for two competing TCP flows. We use the FTP/Generic application provided in GloMoSim, in which a client simply sends data packets to a server without the server sending any control information back to the client other than the acknowledgment packets required by TCP. Whenever a packet is indicated success of delivery by the transport layer (TCP), the client sends the next data packet. The second set shows the results for two competing UDP flows. We use the CBR application in which a client keeps sending data packets to a server at a constant bit rate, such that the sending queue is always non-empty. UDP is the underlying transport layer, thus no acknowledgment packets are sent back to the client. The third set shows the results for one FTP flow competing against one UDP flow, i.e., the FTP/Generic application vs. the CBR application. Except for the difference in the underlying transport layer, both flows generate packets of the same size and they are always backlogged.

We have run each configuration five times with different seed numbers and with a duration of 30 seconds, because we are interested in medium-term fairness in contrast to

101

Table 5.1: Throughput comparison for TCP flows

Conf #	Flow 1	Flow 1 (kbps)	Flow 2	Flow 2 (kbps)	Aggregate (kbps)
2-1	$0 \to 1$	463	$1 \to 0$	469	932
3-1	$0 \to 1$	474	$2 \to 1$	468	942
3-2	$0 \to 1$	452	$1 \to 2$	480	933
3-3	$0 \to 1$	468	$2 \to 1$	404	872
3-4	$0 \to 1$	353	$1 \to 2$	547	899
4-1	$0 \to 1$	0	$2 \to 3$	918	926
4-2*	$1 \to 0$	515±305	$2 \to 3$	419±303	934
4-3*	$0 \to 1$	500±407	$3 \to 2$	406±424	906
4-4	$0 \to 1$	470±78	$2 \to 3$	415±70	884
4-5*	$0 \to 1$	498±67	$3 \to 2$	372±68	870
4-6	$0 \to 1$	475	$2 \to 3$	471	946
4-7	$0 \to 1$	924	$2 \to 3$	0	928
4-8	$0 \to 1$	926	$3 \to 2$	0	928
4-9	$0 \to 1$	427	$2 \to 3$	449	876
4-10*	$0 \to 1$	371	$3 \to 2$	529	901

short-term or long-term fairness.[3] If the standard deviation of throughput is within 10% of the mean throughput, we show mean values only. Otherwise, we show both the mean and the standard deviation of throughput. In these cases, nodes take turns to monopolize the channel for a medium period of time, which we will discuss later.

The results for the three sets are shown in Tables 5.1–5.3. In Tables 5.1 and 5.2, the configurations with results that are worth noting are shown in bold face. If the two competing flows in these configurations are symmetric, they are also shown with asterisks.

We have the following interesting observations from the results shown in Table 5.1 for the two flows that use TCP as the underlying transport layer:

- The flow contention graph that has been used extensively in the past [44, 46–48] is

---

[3]The simulation time is chosen elaborately to expose the medium-term fairness problem.

Table 5.2: Throughput comparison for CBR flows

Conf #	Flow 1	Flow 1 (kbps)	Flow 2	Flow 2 (kbps)	Aggregate (kbps)
2-1	$0 \to 1$	806	$1 \to 0$	799	1600
3-1	$0 \to 1$	806	$2 \to 1$	797	1600
3-2	$0 \to 1$	797	$1 \to 2$	807	1600
3-3*	$0 \to 1$	761±91	$2 \to 1$	782±90	1540
3-4	$0 \to 1$	769	$1 \to 2$	839	1610
4-1	$0 \to 1$	83.4	$2 \to 3$	1500	1580
4-2	$1 \to 0$	820	$2 \to 3$	814	1630
4-3	$0 \to 1$	688	$3 \to 2$	709	1400
4-4	$0 \to 1$	824	$2 \to 3$	808	1630
4-5	$0 \to 1$	725	$3 \to 2$	814	1540
4-6	$0 \to 1$	807	$2 \to 3$	795	1600
4-7	$0 \to 1$	783	$2 \to 3$	824	1600
4-8	$0 \to 1$	1550	$3 \to 2$	28.2	1580
4-9*	$0 \to 1$	734±98	$2 \to 3$	809±94	1540
4-10	$0 \to 1$	781	$3 \to 2$	826	1610

not enough to capture the characteristics of contending flows; we can observe radically different results even though all these two competing flows are the same in the flow contention graph.

- Even for two flows competing in a symmetric way, such as configuration 4-3, the channel bandwidth is not always divided evenly among these two flows. This is the medium-term fairness problem; otherwise, the BEB scheme used in IEEE 802.11 will still be able to achieve fairness in the long run due to the symmetry of the two competing flows.

To illustrate the medium-term problem, we show a snapshot of one simulation run of configuration 4-3 in Fig. 5.4. In this figure, the packet delivery times at MAC layer are recorded and shown. It is clear that the TCP flow between nodes 0 and 1 may monopolize the channel for a very long time. In other simulation runs, it is also possible

Table 5.3: Throughput comparison for competing FTP and CBR flows

Conf #	Flow 1	Flow 1 (kbps)	Flow 2	Flow 2 (kbps)	Aggregate (kbps)
2-1	$0 \rightarrow 1$ (FTP)	0	$1 \rightarrow 0$ (CBR)	1570	1570
3-1	$0 \rightarrow 1$ (FTP)	355	$2 \rightarrow 1$ (CBR)	1000	1360
3-2a	$0 \rightarrow 1$ (FTP)	0	$1 \rightarrow 2$ (CBR)	1570	1570
3-2b	$0 \rightarrow 1$ (CBR)	991	$1 \rightarrow 2$ (FTP)	362	1360
3-3	$0 \rightarrow 1$ (FTP)	268±58	$2 \rightarrow 1$ (CBR)	1110±120	1370
3-4a	$0 \rightarrow 1$ (FTP)	0	$1 \rightarrow 0$ (CBR)	1570	1570
3-4b	$0 \rightarrow 1$ (CBR)	883	$1 \rightarrow 2$ (FTP)	427	1250
4-1a	$0 \rightarrow 1$ (FTP)	0	$2 \rightarrow 3$ (CBR)	1570	1570
4-1b	$0 \rightarrow 1$ (CBR)	102±28	$2 \rightarrow 3$ (FTP)	865	967
4-2	$1 \rightarrow 0$ (FTP)	0	$2 \rightarrow 3$ (CBR)	1570	1570
4-3	$0 \rightarrow 1$ (FTP)	455±93	$3 \rightarrow 2$ (CBR)	815±133	1270
4-4	$0 \rightarrow 1$ (FTP)	340	$2 \rightarrow 3$ (CBR)	983	1320
4-5	$0 \rightarrow 1$ (FTP)	297±65	$3 \rightarrow 2$ (CBR)	1330	1370
4-6	$0 \rightarrow 1$ (FTP)	362	$2 \rightarrow 3$ (CBR)	990	1350
4-7a	$0 \rightarrow 1$ (FTP)	341	$2 \rightarrow 3$ (CBR)	1030	1370
4-7b	$0 \rightarrow 1$ (CBR)	1570	$2 \rightarrow 3$ (FTP)	0	1570
4-8a	$0 \rightarrow 1$ (FTP)	906	$3 \rightarrow 2$ (CBR)	412	948
4-8b	$0 \rightarrow 1$ (CBR)	1570	$1 \rightarrow 2$ (FTP)	0	1570
4-9	$0 \rightarrow 1$ (FTP)	311±50	$2 \rightarrow 3$ (CBR)	1050±110	1360
4-10a	$0 \rightarrow 1$ (FTP)	339	$3 \rightarrow 2$ (CBR)	1030	1370
4-10b	$0 \rightarrow 1$ (CBR)	834±94	$3 \rightarrow 2$ (TCP)	419±55	1250
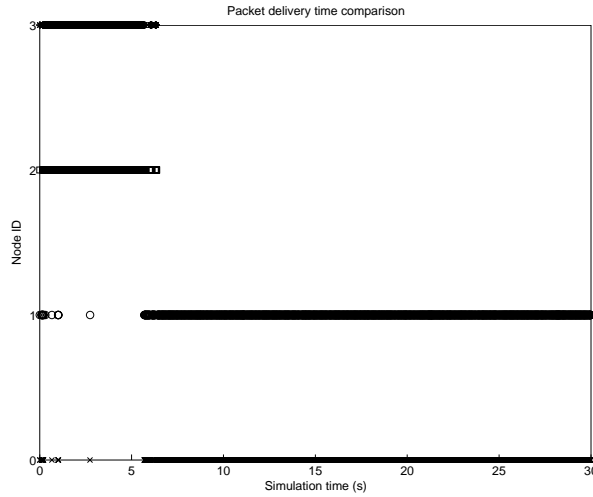
Figure 5.4: A snapshot of two competing TCP flows in conf. 4-3

for the TCP flow between nodes 2 and 3 to monopolize the channel for a very long time.

- In some configurations, such as configurations 4-1, 4-7 and 4-8, the fairness problem is so severe that some TCP connections are in effect prevented from achieving any significant goodput. It should be noted that zero throughput does not mean that TCP connection is not set up. Instead, it is because of the extremely low throughput (on the order of a few kilobytes per second) for these flows that the statistics are not shown in these tables.

For UDP traffic, the fairness problem is not as severe as for TCP traffic. Serious fairness problems occur in only two configurations (configurations 4-1 and 4-8). There are two configurations (configurations 3-3 and 4-9) in which the fairness problem occurs but is not so severe. Some nodes have almost exclusive access to the shared channel for a certain amount of time that is not as long as in the case of TCP traffic. For example, a snapshot of one simulation run of configuration 4-9 is shown in Fig. 5.5. it is clear that the flow from node 0
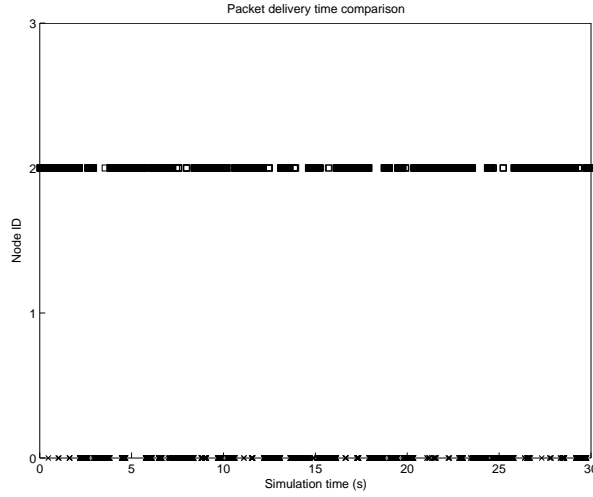
Figure 5.5: A snapshot of two competing UDP flows in conf. 4-9

to 1 may experience very low throughput for several seconds.

When one FTP flow is competing against one CBR flow, as we can expect, the CBR flow achieves much higher throughput than the FTP flow in almost all the cases, except in configurations 4-1b and 4-8a, as shown in Table 5.3. This is because the acknowledgment packets from the FTP server also have to fight their way back to the client in most cases and hence FTP throughput is greatly reduced. Exceptions to that behavior appear in configurations 4-1b and 4-8a, and may be briefly explained as follows. In configuration 4-1b, once node 2 sends a data packet successfully to node 3, node 3 initiates an RTS/CTS handshake for the acknowledgment packet. Because node 1 has already deferred its transmission for the previous successful handshake between nodes 2 and 3, it is very likely that node 0 is still in the backoff stage. Hence, node 2 receives the RTS from node 3 before node 1 receives node 0's RTS, even if node 0 starts sending its RTS during that time. Then, node 2 replies with a CTS, which is received successfully by node 3, while the same CTS may either collide with node 0's transmission at node 1 or be received by node 1. Whatever happens, the CTS from

node 2 forces node 1 to defer its access again. In this way, the flow between nodes 2 and 3 is interrupted only sparsely by the flow from node 0 to node 1 and hence it can achieve much higher throughput than the other flow. The same line of reasoning explains the case for configuration 4-8.

It can also be observed that, in quite a few configurations, the TCP flow is prohibited from being set up or achieving any goodput. In such configurations, we have also experimented with some cases in which the start of CBR traffic is delayed purposely for some time after the start of FTP traffic. We find that the CBR traffic monopolizes the channel sooner or later and then TCP flow fails to achieve any goodput. It is clear that TCP traffic is at a significant disadvantage when competing against UDP traffic due to the acknowledgment packets required by TCP.

We also find that the aggregate throughput of two competing flows is within 10% difference (e.g., configuration 4-5 vs. configuration 4-6) in Table 5.1 and 15% (e.g., configuration 4-3 vs. configuration 4-10) in Table 5.2 when traffic patterns are homogeneous despite the different configurations. This shows that aggregate throughput alone does not reveal the fairness problem easily. However, when traffic patterns are heterogeneous, the fairness problem is more prominent and has negative effects on the aggregate throughput, as shown in Table 5.3, where the difference may be more than 50% (e.g., configuration 4-8a vs. configuration 4-2).

5.5   Conclusion

In this chapter, we have investigated the fairness problem in multi-hop ad hoc networks. We first pinpoint that the required multi-hop coordination can make less effective

those distributed fair queueing schemes that depend on differentiated backoff to prioritize the access of a flow with the minimum service tag. Then we show that the commonly used flow contention graph is insufficient to model the contention among nodes via extensive simulations of two competing flows. Various degrees of the fairness problem can take place due to the different underlying network topologies, despite the same flow contention graph. Our simulation results also reveal that the reverse acknowledgment traffic required by TCP flows has negative effects on both throughput and fairness. On the one hand, TCP acknowledgment traffic can aggravate the fairness problem in the case of homogeneous traffic, because it may reinforce an already leading flow in some cases such that the flow can gain exclusive access to the shared channel for a long time. On the other hand, a TCP flow is at a disadvantage in competing against UDP flows, because TCP acknowledgment traffic has to fight its way back to the source, and throughput can be degraded in most cases due to the interference from unregulated UDP flows. This motivates our work to be presented in Chapter 6 in which we propose a novel hybrid channel access scheme having better throughput and fairness properties and in Chapter 7 in which we propose framework and mechanisms for fair medium access.

Chapter 6

A Novel Hybrid Channel Access

Scheme

6.1   Introduction

In Chapters 2 and 3, we have investigated some sender-initiated MAC schemes that use omni-directional and directional antennas. There is also another category of MAC schemes that are receiver-initiated, in which a node polls its neighbors actively to see if they have packets for itself. The rationale behind receiver-initiated schemes is that, a receiver usually has better knowledge of the contention around itself and collision avoidance is more important at the receiver's side as the receiver needs to receive relatively long data packets successfully which are more vulnerable to interference. It has been shown that, if the polled nodes always have packets for the polling node, receiver-initiated schemes with proper colli-sion avoidance procedures can outperform sender-initiated schemes by reducing the overhead

of control packets [23, 24]. Otherwise, the performance may degrade due to wasted transmissions of polling packets that poll inactive nodes with no packets for the polling node. The degradation in performance will be more conspicuous in light to medium traffic load, unless a good traffic predictor is available at the polling node.

Despite the potential benefits of receiver-initiated schemes, they have not received wide acceptance. One reason is that sender-initiated schemes are more straightforward, because a sender has full knowledge of the packets in its queue and it can initiate the collision avoidance handshake only when necessary. On the other hand, for receiver-initiated schemes, a good traffic estimator and an appropriate polling discipline that can be adapted to the dynamic environments of ad hoc networks are mandatory and they have not been investigated sufficiently so far. Another reason is the prevalent acceptance of the IEEE 802.11 MAC protocol in the research community, which uses a sender-initiated collision avoidance scheme. Many performance enhancements have been proposed and they are confined to the sender-initiated framework stipulated by the IEEE 802.11 MAC protocol.

Despite its popularity, the IEEE 802.11 MAC protocol can suffer severe fairness problems in multi-hop ad hoc networks where location-dependent contention is common and the binary exponential backoff (BEB) scheme aggravates the fairness problem. Although two basic classes of schemes have been proposed in the recent past to address the fairness problem in IEEE 802.11, they are insufficient as discussed in Chapter 5. Besides, despite the differences of backoff algorithms and information exchange among these schemes, the underlying channel access scheme remains largely the basic sender-initiated collision avoidance handshake, which can be less effective than a receiver-initiated scheme when a receiver has

better knowledge of the contention around itself than the sender. This motivates us to design an adaptive collision avoidance scheme that makes use of both sender-initiated and receiver-initiated handshakes, because a receiver-initiated handshake is more desirable in some cases and a better tradeoff between throughput and fairness may be achieved. To expedite its introduction, the new hybrid scheme should fit within the IEEE 802.11 framework, even though it combines both sender-initiated and receiver-initiated handshake, and nodes implementing the new scheme should not break an existing 802.11-based network. Furthermore, the new scheme should be simple and not introduce new types of control packets, because they may complicate implementation of the finite state machine of the protocol and degrade the overall network throughput unnecessarily when the basic sender-initiated scheme suffices.

The rest of the chapter is organized as follows. In Section 6.2, the new hybrid scheme is specified, which in fact is a very simple extension to the existing IEEE 802.11 MAC protocol and involves only some additional queue management and book-keeping work. In Section 6.3, a measurement-based fair scheme [49] is described, which is one of the simple and straightforward fair schemes proposed so far; it does not require explicit information exchange among nodes and serves as a basis for comparison with our new hybrid scheme. In Section 6.4, simulations with the original IEEE 802.11 MAC protocol, the measurement-based fair scheme and the new hybrid scheme are presented for both UDP- and TCP-based traffic. It is shown that although the new hybrid scheme cannot solve the fairness problem conclusively, it can alleviate the fairness problem in some cases with almost no degradation in throughput. It is also reasoned that more explicit information exchange among nodes is mandatory to solve the fairness problem while maintaining reasonable throughput. Section 6.5 concludes this

111

chapter.

6.2   The New Hybrid Collision Avoidance Scheme

Bharghavan et al. [14] proposed a request-for-request-to-send (RRTS) packet to alleviate some of the interference problems due to hidden terminals in their seminal paper to address the fairness problem. Talucci and Gerla [18] proposed MACA-BI (Multiple Access with Collision Avoidance - By Invitation) which was the first receiver-initiated MAC protocol. Garcia-Luna-Aceves and Tzamaloukas [24] advanced that work and proposed several collision-free RIMA (receiver-initiated multiple access) protocols. Here collision-free means that, once a node sends a data packet, the data packet can be received by the receiver successfully, given that the channel is ideal without impairment and the only cause of failure to receive a packet is concurrent transmissions from multiple nodes. RIMA protocols achieve this collision-free property by introducing some additional types of short control packets and enforcing various collision-avoidance waiting periods. The receiver-initiated handshake in our proposed hybrid channel access scheme is simpler than that in the RIMA protocols. Firstly, it does not introduce new types control packets. Instead, a CTS packet is used as the polling packet to maintain compatibility with the original IEEE 802.11 MAC protocol. Secondly, it does not include the various collision-avoidance waiting periods enforced in RIMA protocols. Instead, nodes defer access to the shared channel according to the network allocation vector (NAV) included in those overheard packets, which specifies the duration of the ensuing handshake. The reason is that the IEEE 802.11 MAC protocol itself cannot ensure collision-free data packet transmissions. We opt not to introduce additional collision-avoidance procedures

and try to maintain compatibility with the existing protocols. Hence, the receiver-initiated collision avoidance handshake just includes a three-way CTS-data-ACK exchange between polling and polled nodes. Though it is not expected that the hybrid scheme will improve throughput because it does not provide strict collision avoidance, it may still alleviate the fairness problem, because both a sender and a receiver can initiate a collision avoidance handshake alternately and the burden of contending for the shared channel is distributed to participating nodes according to the different degrees of contention they experience.

Our hybrid collision avoidance scheme is built around the framework of the IEEE 802.11 MAC protocol. A node that implements this scheme operates alternately in two modes, sender-initiated (SI) and receive-initiated (RI). The SI mode is the default mode, which is in effect the same as the original IEEE 802.11 MAC protocol. The usual four-way RTS-CTS-data-ACK handshake is used in the SI mode. The aforementioned receiver-initiated three-way collision avoidance handshake is used in the RI mode introduced in the hybrid scheme, and is triggered only when the SI mode does not perform well. In this mode, more cooperation between a pair of sending and receiving nodes is required, because both of them need to enter the RI mode before the receiver-initiated handshake can be initiated.

The only necessary change to the frame structures in the IEEE 802.11 standard to implement the hybrid scheme is the addition of the RI flag. Figure 6.1 illustrates the frame structure of the IEEE 802.11 RTS frame (ref. Fig. 13 in Page 35 and Fig. 16 in Page 41 of the IEEE 802.11 standard [16]). Given that the *More data* bit is not used in the ad hoc mode according to the standard, it may be reused as the RI flag to indicate if the RI mode is on or not. Nodes that do not implement the hybrid collision avoidance scheme can safely ignore this

113
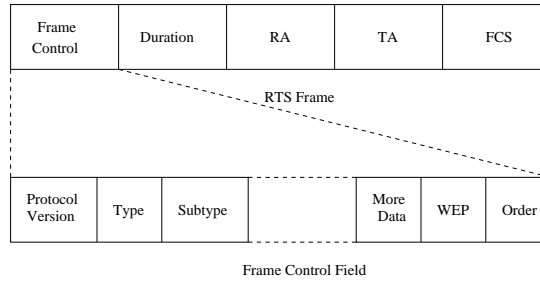
Figure 6.1: Illustration of the IEEE 802.11 frame structure

bit. The states of both sending and receiving nodes in the proposed hybrid collision avoidance scheme are shown in Figure 6.2 and are explained separately. A sender enters *RI setup* mode when it sends the same RTS packet for more than one half of the times allowed in the IEEE 802.11 MAC protocol and has no response from the intended receiver. Failure to obtain a response from the intended receiver usually implies that contention around the receiver is so severe that the receiver is prevented from responding. Hence, it is more appropriate to let the receiver start the collision-avoidance handshake when this happens. The number of unsuccessful RTS packets to activate the RI mode is chosen according to the simulations with the network configurations investigated in this chapter. The present value in the scheme is able to achieve balanced results while other values may perform very differently for different configurations.

After the sender enters *RI setup* mode, it sets the RI flag in all the subsequent RTS packets and other packets that it sends out and requests the intended receiver to enter the RI mode as well. During this stage, the node keeps sending RTS packets following the usual collision-avoidance procedures, because it has not established an *association* with the intended receiver. There are two possible outcomes. One outcome is that the node never gets any CTS packet from the intended receiver. In this case, the sender may declare the receiver down after

114
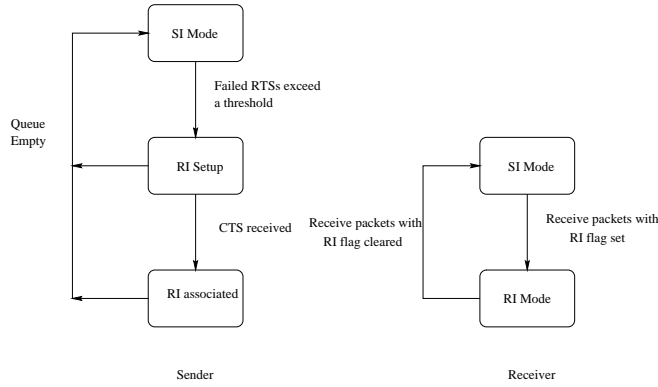
Figure 6.2: State transition diagram of sending and receiving nodes

it has to drop a few packets. The other outcome is that it receives a CTS packet from the intended receiver. In this case, the sender enters the *RI-associated* mode and will not send an RTS to the receiver thereafter. This helps to reduce the contention around the receiver and also makes the sender available for accepting polling requests from the receiver. To keep the receiver in the RI mode, the sender keeps setting the RI flag in all the data packets that it sends out. The RI flag is cleared only when the sender's queue becomes empty.

The receiver enters and stays in the RI mode when it receives RTS packets or data packets destined to it with the RI flag set. The receiver then generates RI-response packets (which are in fact self-initiated CTS packets) and multiplexes them with other data packets in its MAC queue. However, the receiver should not generate RI-response packets indiscriminately when it receives a packet with the RI flag on, lest serious fairness problem may occur. This can be explained as follows. When an RI-response packet becomes the head-of-line (HOL) packet of a receiver's queue, the node will send a self-initiated CTS to the sender, which in fact serves as the ready-to-receive (RTR) packet to poll the sender in the RIMA protocols [24]. If the sender replies with a data packet with the RI flag still on, which implies that

there are more packets in its sending queue, the receiver will add another RI-response packet to the end of its queue. If there is no packet for other nodes intervened in the MAC queue, the receiver will be *locked into* the sender and will keep sending CTS packets to it. In this way, they may monopolize the shared channel for a long time, which obviously defeats the purpose of the hybrid scheme. Hence, when a node receives a packet with the RI flag on, it checks its HOL packet to see whether it is an RI-response packet for the node that just sent this packet. If so, the RI request is ignored; otherwise, it is added to the end of its MAC queue.

The RI-response packets are treated like RTS packets for normal data packets. That is, when they are served via a successful receiver-initiated CTS-data-ACK handshake or when they are transmitted more than the times allowed for RTS packets in the IEEE 802.11 standard, they will be removed from the MAC queue. Such precautions are necessary. One reason is to avoid excessive delay or deadlock when the sending node is down or moves out of range of the receiving nodes. Another reason is to promote fairness so that neighboring nodes may still get chances to initiate handshake with the receiver or other nodes.

The above specification clearly shows that, with some additional queue management and book-keeping work, the existing IEEE 802.11 can be easily extended to support a receiver-initiated scheme while maintaining compatibility.

6.3   The Measurement-based Fair Scheme

In this section, we describe the measurement-based fair scheme [49] with which we compare the IEEE 802.11 MAC protocol and the new hybrid scheme. The rationale behind the scheme is surprisingly simple. Whenever a node sends or receives a packet, it updates its

own estimation of its share (W_{ei}) or other nodes' share (W_{eo}) of the channel depending on the purpose of the packet. To avoid any explicit information exchange among these nodes, each node just treats all the nodes around itself as a single entity that competes against itself. For example, if a node sends an RTS packet, it will update W_{ei} because the RTS packet serves to reserve the channel for itself. If the node receives an RTS packet addressed to itself, it updates W_{eo} because the RTS packet serves for other nodes. Details on the updating of W_{ei} and W_{eo} can be found in [49] and are not repeated here. Then the ratio between W_{ei} and W_{eo}, which is denoted by FI_e, serves as a fairness index to show whether a node is leading or lagging in channel access. If FI_e for a node is larger than a pre-defined constant C ($C > 1$), which implies that the node has obtained more than its fair share, the node doubles its contention window (CW) from which the backoff timer is derived. If FI_e lies between $1/C$ and C, then the node just holds on to its current CW as it estimates that both its neighbors and itself have obtained roughly equal shares. If FI_e is smaller than $1/C$, then the node cuts its CW to a half to contend more vigorously for the channel. It should be noted that CW is bounded by the minimum and maximum values stipulated in the IEEE 802.11 standard. The measurement-based fair scheme is shown to be quite effective in the configurations investigated in [49] by sacrificing some throughput for better fairness.

However, this scheme may encounter the problem of severe throughput degradation in some cases, e.g., when two neighboring nodes are engaged in TCP-based connections. This can be explained as follows. In the measurement-based scheme, a node at one end of a TCP connection continuously estimates its share and other node's share of the channel including the node on the other side of the connection. When this node sends one or a few data

117

packets, it estimates that its use of the channel has exceeded its fair share and will increase

its contention window accordingly. The node at the other end of the TCP connection behaves

similarly. In this way, both nodes may have a contention window that is larger than necessary

and the throughput is degraded due to the increased time wasted in waiting. The degradation

in throughput can also happen in UDP-based traffic in which two nodes take turns in channel

access according to the their own measurements. However, this phenomenon can be more

conspicuous in TCP-based connections, because flow control in TCP may also be activated,

which can further slow down the channel access activities unnecessarily.

6.4   Simulation Results

In our simulations, we focus on how two competing flows share the available channel resource in a few simple network configurations. These configurations are already shown previously in Figures 5.2 and 5.3.

We use the same simulation settings as used in Chapter 5 and investigate the performance of the IEEE 802.11 MAC protocol, the measurement-based fair scheme (MFS) and the new hybrid scheme under both UDP- and TCP-based traffic. In the first set of simulation experiments, there are two competing UDP-based flows. Tables 5.2 in Section 5.4 have shown the performance of the original IEEE 802.11 MAC protocol. For configurations 4-1 and 4-8, some nodes are almost denied access to the shared channel and suffer severe degradation in throughput. For other configurations, the original MAC protocol works fairly well. Table 6.1 shows the performance of the original IEEE 802.11 MAC protocol, the MFS scheme and the hybrid scheme. It is apparent that, the hybrid scheme is the same as as the original IEEE

802.11 when the RI mode is not triggered in some network configurations. For simplicity, the performance of all schemes in these configurations is not shown here, because the conclusions to be drawn are unaffected. From Table 6.1, it is clear that the fairness problems in configurations 4-1 and 4-8 are alleviated significantly, without a sacrifice in throughput when the hybrid scheme is used. In addition, even when the RI mode is triggered unnecessarily in the three other configurations, it has almost no negative effect on throughput. On the other hand, the MFS works best in configurations 4-1 and 4-8. However, in other configurations in which the original IEEE 802.11 MAC protocol works well, the use of MFS degrades throughput unnecessarily, in some cases to less than 50% of the original. It is also worth noting that the aggregate throughput in all these network configurations remains almost the same despite the fact that the fairness problem exists in some configurations. This shows the importance of considering the underlying topologies of networks even though they have the same flow contention graph.

In the second set of simulation experiments, there are two competing TCP-based flows. We use the FTP/Generic application provided in GloMoSim which has been described in detail in Section 5.4. It should be noted that the acknowledgment packet from TCP is still regarded as a normal data packet from the view of MAC layer. Hence, due to the peculiarities of the application, it is disadvantageous for the MAC layer to transmit more than one packet at a time. When this is applied to the hybrid scheme, it means that it is more desirable for a node and its peer to leave RI associated mode just after a CTS-Data-ACK handshake is done, so that they can switch the roles of sender and receiver timely. So in the implementation of the hybrid scheme, we make the necessary changes to take this into account. That is, for such type

Table 6.1: Throughput comparison for the IEEE 802.11, the measurement-based fair scheme (MFS) and the hybrid scheme (with RI mode) – two CBR flows (throughput measured in kbps)

Config #	Scheme	Flow #	Throughput	Flow #	Throughput	Aggregate
3-3	802.11	$0 \rightarrow 1$	761	$2 \rightarrow 1$	783	1540
	+RImode	$0 \rightarrow 1$	794	$2 \rightarrow 1$	774	1610
	+MFS	$0 \rightarrow 1$	472	$2 \rightarrow 1$	471	943
4-1	802.11	$0 \rightarrow 1$	83.4	$2 \rightarrow 3$	1500	1580
	+RImode	$0 \rightarrow 1$	36.9	$2 \rightarrow 3$	1230	1600
	+MFS	$0 \rightarrow 1$	979	$2 \rightarrow 3$	534	1510
4-3	802.11	$0 \rightarrow 1$	688	$3 \rightarrow 2$	709	1400
	+RImode	$0 \rightarrow 1$	665	$3 \rightarrow 2$	643	1310
	+MFS	$0 \rightarrow 1$	691	$3 \rightarrow 2$	698	1390
4-8	802.11	$0 \rightarrow 1$	1550	$3 \rightarrow 2$	28.1	1580
	+RImode	$0 \rightarrow 1$	1280	$3 \rightarrow 2$	319	1600
	+MFS	$0 \rightarrow 1$	522	$3 \rightarrow 2$	986	1510
4-9	802.11	$0 \rightarrow 1$	734	$2 \rightarrow 3$	809	1540
	+RImode	$0 \rightarrow 1$	815	$2 \rightarrow 3$	742	1560
	+MFS	$0 \rightarrow 1$	472	$2 \rightarrow 3$	471	942

of traffic, a node clears its RI flag when it receives acknowledgment packet from the node that

it is sending data packets to. Table 5.1 in Section 5.4 shows the performance of the original

IEEE 802.11 MAC protocol and Table 6.2 compares it with the MFS and the hybrid scheme.

It is clear that the fairness problem is much more severe for two competing TCP-based flows

than for the case of UDP-based flows. In some cases, such as configurations 4-1, 4-7 and 4-8,

one FTP flow is denied access to the shared channel for most of the time. When the hybrid

collision avoidance scheme is used, in some cases it is triggered and performs almost the

same as the original IEEE 802.11 MAC protocol while in some other cases it is not triggered

at all. For simplicity Table 6.2 shows only the results when there exist differences between

these two schemes. It is clear that the hybrid scheme performs slightly better than the original

802.11 MAC scheme for configurations 3-3 and 4-2, while it performs much better in terms

Table 6.2: Throughput comparison for the IEEE 802.11, the measurement-based fair scheme (MFS) and the hybrid scheme (with RI mode) – two FTP flows (throughput measured in kbps)

Config #	Scheme	Flow #	Throughput	Flow #	Throughput	Aggregate
3-4	802.11	$0 \to 1$	352	$1 \to 2$	548	900
	+RImode	$0 \to 1$	330	$1 \to 2$	554	885
	+MFS	$0 \to 1$	149	$1 \to 2$	176	325
4-1	802.11	$0 \to 1$	0	$2 \to 3$	926	929
	+RImode	$0 \to 1$	-	$2 \to 3$	-	-
	+MFS	$0 \to 1$	216	$2 \to 3$	228	444
4-2	802.11	$1 \to 0$	488±103	$2 \to 3$	453±102	942
	+RImode	$0 \to 1$	439±99	$3 \to 2$	502±98	940
	+MFS	$0 \to 1$	249	$3 \to 2$	252	502
4-3	802.11	$0 \to 1$	530±432	$3 \to 2$	392±438	922
	+RImode	$0 \to 1$	397±71	$3 \to 2$	455±78	852
	+MFS	$0 \to 1$	220	$3 \to 2$	220	441
4-7	802.11	$0 \to 1$	928	$2 \to 3$	0	930
	+RImode	$0 \to 1$	-	$2 \to 3$	-	-
	+MFS	$0 \to 1$	424	$2 \to 3$	404	828
4-8	802.11	$0 \to 1$	929	$3 \to 2$	0	930
	+RImode	$0 \to 1$	-	$3 \to 2$	-	-
	+MFS	$0 \to 1$	246	$3 \to 2$	206	452

of fairness in configuration 4-3, though there is about an 8% degradation in throughput. It is more difficult to improve fairness of TCP-based flows than UDP-based flows due to the flow control and congestion avoidance functions of TCP. A node that suffers excessive packet loss or delay decreases its sending rate according to TCP, which can aggravate the fairness problem already existing at the MAC layer. In such cases, even the hybrid scheme can lose its effectiveness.

On the other hand, the MFS achieves very good fairness in all these configurations but at the cost of much reduced throughput except for configuration 4-7. This is due to the fact that nodes are slowed down to encourage fair contention for the shared channel.

From the radically different results of these networks that share the same flow con-

tention graph, it can be reasoned that the proposed fair schemes that are based on flow contention graphs (e.g., [46–48]) are not sufficient to solve the fairness problem conclusively. Simple schemes like the MFS cannot solve the fairness problem desirably as well if reasonable throughput is to be maintained. Additionally, though the new hybrid scheme has better throughput and fairness properties, it is also inadequate to solve the fairness problem because it does not make any modification to the existing (unfair) backoff algorithm in the IEEE 802.11 MAC protocol and does not rely on additional contention information, which may be useful for nodes to contend for the shared channel more efficiently and fairly. Hence, more explicit information exchange among nodes as well as the good use of such information should be studied to address both fairness and throughput adequately.

6.5 Conclusion

In this chapter, we have proposed a new hybrid channel access scheme that includes sender-initiated and receiver-initiated collision avoidance. This is based on the observation that sometimes a receiver-initiated scheme is more appropriate when receivers are more knowledgeable of the contention around themselves and can compete for the channel more effectively. By adaptively sharing the burden of initiating the collision-avoidance handshake between the nodes that experience different levels of contention, better fairness may be achieved with almost no degradation in throughput. An attractive feature of the new scheme is that it is a simple extension to the existing IEEE 802.11 MAC protocol and maintains compatibility with the standard. Simulations are conducted with the original IEEE 802.11 MAC protocol, a measurement-based fair scheme (MFS), and the new scheme. It is shown that, al-

though the proposed hybrid scheme does not solve the fairness problem conclusively, it does alleviate the fairness problem in some cases without sacrificing much throughput and simplicity. Simple schemes such as the MFS can achieve far better fairness but sacrifice too much throughput. Hence, it is reasoned that more explicit information exchange among nodes is needed to solve the fairness problem conclusively and this leads to our work to be presented in Chapter 7 which addresses the problem in a more systematic way.

Chapter 7

Framework and Mechanisms for Fair Medium Access

7.1 Introduction

In Chapter 6, we propose a hybrid channel access scheme that combines both sender-initiated and receiver-initiated collision avoidance handshake to alleviate the fairness problem. The attractiveness of this approach is that it is compatible with the IEEE 802.11 framework and involves only some additional queue management and book-keeping work. However, we have also shown that, despite its simplicity, it is not very effective for TCP-based flows and that more information exchange among nodes is necessary to solve the fairness problem conclusively. This motivates us to propose the framework, which is presented in the rest of this chapter, to address the fairness problem in a systematic way. In Section 7.2, we identify several key components that constitute our fairness framework and explain the rationale for their

necessity. In Section 7.3, we propose new algorithms to realize the fairness framework. The resulting scheme, which we simply call topology aware fair access (TAFA) is evaluated in Section 7.4 through computer simulations. The performance of TAFA is compared with that of the original IEEE 802.11 MAC protocol and the hybrid channel access scheme proposed in Chapter 6 for both UDP- and TCP-based traffic. Simulation results show that TAFA can solve the fairness problem in UDP-based applications with negligible degradation in throughput. It can also solve the notorious problem of starvation of flows in TCP-based applications, despite some moderate degradation in throughput. Hence, TAFA shows a much better overall tradeoff between throughput and fairness than the other schemes investigated. Section 7.5 concludes this chapter with directions for future work.

7.2   The Fairness Framework

In this section, we describe a framework for achieving better fairness consisting of four key components: Exchanging flow information among nodes, using an adaptive backoff algorithm that is as stable as binary exponential backoff (BEB) but does not have the inherent deficiency of aggravating the fairness problem, switching sender-initiated and receiver-initiated scheme as appropriate, and dealing with two-way flows.

The need for the exchange and maintenance of flow-contention information can be illustrated by a simple example with the network configuration 4-8 shown in Figure 5.3. For configuration 4-8, node 2 knows that both node 0 and node 3 are sending nodes. However, if node 2 does not explicitly tell both node 0 and node 3 about the existence of each other, the handshake between node 0 and node 1 will tend to dominate the channel, because node 3's

transmissions will mostly collide with either node 0 or node 1's transmissions at node 2, and both node 0 and node 1 may incorrectly perceive that node 0 and node 1 are the only active nodes in the network. Even though they may receive node 2's packets sporadically and make some ad hoc adjustment, without a systematic way to obtain flow information, the fairness problem cannot be solved conclusively.

The second component of our framework is an adaptive backoff scheme that is mandatory because the existing binary exponential backoff can aggravate the fairness problem, as shown extensively in the literature [14, 42, 84] and Chapter 5. Nodes should decide their channel access based on the information of competing flows gathered through the first component.

The third component of our framework is a hybrid channel access scheme that combines both sender-initiated and receiver-initiated collision handshake. This is largely due to the advantage of distributing the burden of initiating the collision avoidance handshake between a pair of sending and receiving nodes depending on the different degrees of contention they experience. For example, in the network configuration 4-1 shown in Figure 5.3, the flow from node 0 to node 1 will suffer severe throughput degradation if no proper action is taken, because RTS from node 2 can always be received by node 3 successfully while node 0's RTS collides with node 2's transmissions at node 1 most of the time. In this case, if the collision avoidance is initiated by node 1, which transmits CTS to node 0 directly, then the channel bandwidth will be shared between these two flows more evenly, because node 1 and node 2 are direct neighbors and it is easier for them to coordinate their access to the channel.

The fourth component of our framework is a key contribution of the framework and

consists of dealing with two-way traffic in which there are one data flow and one acknowledgment flow between two nodes, as is the case in most TCP-based flows. In such cases, usually one node cannot continue sending data packets, unless it receives upper level acknowledgment packets from the other node. Though viewed from a traditional MAC's perspective they are separate flows, the performance of these two flows is coupled and they should compete as a collective entity rather than do so separately. Fairness for such cases is only touched upon in Chapter 6 and has not been addressed adequately in the literature, because most of the performance evaluation of fair MAC schemes so far has been done with constant bit rate (CBR) like traffic. The information about whether a flow is one-way or has a reverse flow can be conveyed from the application layer down to the MAC layer through some interface. For example, one bit in the protocol headers from upper layers can indicate the presence or absence of an acknowledgment flow. Though the details need to be worked out in a cross-layer approach and are beyond the scope of the thesis work, we believe that such information and hence the required special processing are necessary to achieve the desired fairness goal.

7.3 Topology Aware Fair Access

The topology aware fair access (TAFA) scheme is a realization of the fairness framework described previously, and consists of four parts corresponding to the four components in the framework.

7.3.1 Exchange and maintenance of flow information

Each node maintains a flow table and each entry in the table contains the following information about a flow: *source address*, *destination address*, *service tag*, *direct flag* and *position flag*.

The service tag is used to measure how much channel resource the flow has received. Though there can be several ways to calculate the service tag, we use a simple one, which consists of the number of bytes that have been sent by the sender and acknowledged by the receiver. The service tag is updated by the sender when it receives an acknowledgment from the receiver and updated information is propagated to other nodes through subsequent packet transmissions.

The direct flag is used to indicate whether the flow is known directly through listening to the channel or indirectly through flow advertisements from other nodes. For example, in the network configuration 4-8 shown in Figure 5.3, node 3 cannot know the flow from node 0 to node 1 directly and has to rely on node 2 to advertise that flow to it. In this case, the flow from node 0 to node 1 is recorded as indirect in node 3's flow table and node 3 does not advertise the indirect flow.

The position flag is used to indicate whether a flow is original, a derivative, or not applicable to either case. This flag is used to handle two-way traffic. For example, in some TCP-based applications, one end of the connection cannot continue sending packets, unless it receives a TCP acknowledgment from the other end. The MAC protocol cannot just treat the data flow and the acknowledgment flow as separate flows. Due to the asymmetry of most connections, i.e., a data flow usually generates much more traffic than the corresponding

acknowledgment flow, trying to equate the channel utilization for both flows would lead to throughput degradation. So it is important to use the position flag to indicate whether the flow is *original* (data flow) or *derivative* (acknowledgment flow) and the service tag of a derivative flow should be adjusted according to that of the corresponding original flow.

In this scheme, an RTS or a CTS only carry the information about the current flow (from the sender to the receiver) to reduce the fixed overhead that exists whether fairness is desired or not. Because the source and destination of a flow are self evident and a direct flag is not necessary, the extra information included in the RTS and CTS is just the service tag and the position flag of the flow. A receiver just copies the service tag in an RTS to its outgoing CTS, so that the neighbors of the receiver can also know the service tag of the ongoing flow. On the other hand, data packets and ACKs carry extra information about other flows maintained by the node if necessary. The rationale for treating these control packets differently is that the size of an RTS and a CTS can be fixed and nodes can get the duration information of the subsequent handshake from the network allocation vector (NAV) embedded in all packets. Because data packets are of varying size, it is acceptable for them to carry a bit more information. An ACK should also carry some extra flow information, otherwise those nodes that are neighbors of the node sending the ACK will never get any information about the flows around the node if the node does not send any data packet.

Specifically, to reduce the overhead incurred in the flow information exchange, nodes advertise only one flow at a time in the data or ACK packets they transmit, and one flow is chosen from the node's flow table in a round-robin way. As stated earlier, they only advertise flows that they know directly through receiving transmissions from either the sender

129

or the receiver of the flow, rather than through the advertisement by other nodes. In this way, a node only gets the information about those flows that are one hop and two hops away from itself. This avoids building up all the flows' information in a node which is unnecessary because channel access should be a local decision based only on the information of flows competing directly to avoid the complexity of making global decisions which is not what MAC layer should consider. Besides, nodes can obtain the updates of neighbor flows more quickly because only such flows are advertised. Flow information adversed in data and ACK packets includes only the source address, destination address and service tag.

Through the advertisement of flows, a node comes to know the other flows that may be competing with itself, gathers neighborhood topology information gradually, and adjusts its channel access accordingly.

7.3.2 Flow aware backoff algorithm

The flow aware backoff algorithm runs whenever a new contention window needs to be calculated, for example, after the first transmission of an RTS packet fails to receive CTS in due time or after the successful reception of an ACK packet. In this algorithm, each node also maintains two flags: *MyFlow* and *OtherFlow*. These two flags are used for a node to decide its contention window (CW), which is the upper bound of the uniform distribution from which a backoff timer is generated.

When a node receives the acknowledgment for its data packet, it updates its service tag and sets *MyFlow* true. When a node receives updated and greater service tag for other flows, it sets *OtherFlow* true. If a node receives a smaller service tag for other flows than that

recorded in its table, it does not change *OtherFlow* flag because this is obsolete information from other nodes that have not got the up-to-date information.

Unlike other algorithms that deviate significantly from the binary exponential back-off (BEB) used in the IEEE 802.11 MAC protocol, we adopt BEB's basic idea of quick contention resolution and robustness and the resulting backoff algorithm is shown in Figure 7.1 in pseudo-code. Lines from 1 through 7 deal with the case when the node is the sender of the flow with the minimum service tag. If neither the flow nor any other flow progresses (lines 2–3), then it means that some other nodes may also perceive that they have the minimum flows and it is important for the node to double its contention window (CW) for quick contention resolution. If any other flow progresses (lines 4–5), then the node should keep its current CW lest it may cause collisions by decreasing the CW and suffer unfairness by increasing the current CW, because it is already lagging behind other flows. If this flow has already made progress (lines 6–7), then it is safe to set its CW to the minimum value, because there is no perceived immediate contention from other flows. Lines from 9 through 17 deal with the case when the node does not have the minimum flow. If neither my flow nor other flow progresses (lines 9–10), it is important to double the CW for quick contention resolution. If only other flows make progress (lines 11–12), then it is adequate to keep the current CW, because the node does not require immediate access to the channel. However, if only my flow progresses (lines 13-14), then it means that the node is too aggressive in its channel access and should double its CW to yield the channel access to the other nodes that have the minimum flows. If both my flow and other flow progress, then the node can reset the CW to the minimum value to avoid too much time spent in backoff. At last, in line 18, both *MyFlow* and *OtherFlow* are

```
1: if (My flow has the min service tag in my flow table) {
2:      if (!MyFlow && !OtherFlow)
3:             Double contention window;
4:      else if (OtherFlow)
5:             Keep current contention window;
6:      else if (MyFlow)
7:             Reset contention window to minimum;
8: } else {
9:      if (!MyFlow && !OtherFlow)
10:            Double contention window;
11:     else if (OtherFlow && !MyFlow)
12:            Keep current contention window;
13:     else if (MyFlow && !OtherFlow)
14:            Double contention window;
15:     else if (MyFlow && OtherFlow)
16:            Reset contention window to minimum;
17: }
18: Clear MyFlow and OtherFlow.
```

Figure 7.1: The adaptive backoff algorithm

cleared and the backoff algorithm will be adapted again to any future change made to these two flags.

It should be noted that in the backoff algorithm, nodes adjust their contention window according to not just the flow with the minimum service tag, but also the progress status of its own flows and other flows, so it is very important for nodes to advertise all the flows they know to their neighbors and then they can adapt the changes of flows around them.

7.3.3 Topology-aware hybrid collision avoidance handshake

As we have discussed, sometimes receiver-initiated collision avoidance can be more effective than sender-initiated and a combination of both is shown to yield quite satisfactory results when used to address the fairness problem as discussed in Chapter 6.

Remember that, in the hybrid channel access scheme, there are both sender-initiated and receiver-initiated handshakes and the criterion to trigger the receiver-initiated handshake is that a node sets the RI request flag in its packets after it has sent the same RTS packet for more than one half of the times allowed in the IEEE 802.11 MAC protocol and receives no response from the intended receiver.

The problem with this approach is that the receiver can hardly get any RTS sometimes due to high contention around it and hence receiver-initiated handshake cannot be triggered. This phenomenon is especially conspicuous for a two-way TCP connection, which consists of one data flow and one acknowledgment flow, because a pair of nodes may take turns to grab the channel, while other less privileged nodes may defer their access to the channel further due to the flow control and congestion avoidance functions in TCP.

To address the above problem, we propose a topology-aware scheme that switches between sender-initiated and receiver-initiated handshake. The basic idea is to make nodes that are closer to the contention initiate the handshake. To facilitate the description of the algorithm, some notations are introduced and shown in Table 7.1. Two flows are called dependent if they need to take turns to proceed, like a data flow and an acknowledgment flow in most TCP-based flows. That is why the position flag is exchanged and recorded in a node's flow table.

Figure 7.2 shows the criteria to switch between sender-initiated and receiver-initiated handshake. Similar to the algorithm shown in Figure 7.1, lines from 1 through 7 deal with the case when the node is the sender of the flow with the minimum service tag. If there is any independent flow in this node's table (lines 2–6), then the node needs to differentiate between

133

Table 7.1: Notations used in the hybrid scheme

V	The node applying the algorithm
f_m	The flow with the minimum service tag among all the flows in node V's flow table
f_{mi}	The flow with the minimum service tag among all the flows in node V's flow table that are not dependent on flows originating from node V
$S(f)$	Sender of flow f
$R(f)$	Receiver of flow f
$N(V)$	Node V's neighbors

two cases. If either the sender or the receiver of the independent flow which has the minimum service tag is this node's neighbor (line 3), then the usual sender-initiated handshake is used (line 4). Otherwise, it is possible that the receiver of the node is closer to either the sender or the receiver of that independent flow and it is more appropriate for the node to ask its receiver to use receiver-initiated handshake (line 6). In this way, the node and its receiver may compete for the channel more effectively. If the node does not have the minimum flow (lines 8–13), it should find the minimum flow in its flow table first. If the node is the receiver of the minimum flow or either the sender or the receiver of the minimum flow is its neighbor, then it just stays in the SI mode (lines 9–11). Otherwise, it means that the receiver of its flow may be closer to the nodes having the minimum flow, and then the node asks its receiver to enter the RI mode (line 12) with the hope that its receiver may compete for the channel more effectively than itself.

```
1:   if (S(f_m) is V) { // Our flow has the min service tag
2:         if (∃f_mi) { // If an independent flow is found;
3:               if (S(f_mi) ∈ N(V) or R(f_mi) ∈ N(V))
4:                     Sender-initiated;
5:               else
6:                     Receiver-initiated;
7:         } else Sender-initiated;
8:   } else { // Some other flow has the min service tag
9:         if (V is R(f_m) or S(f_m) ∈ N(V)
10:                    or R(f_m) ∈ N(V))
11:               Sender-initiated;
12:         else   Receiver-initiated;
13: }
```

Figure 7.2: The criteria to choose sender-initiated or receiver-initiated handshake

7.3.4   Dealing with two-way flows

Two-way flows require special processing as discussed before. We describe some necessary changes to the algorithms discussed in the previous subsections.

For an original flow and a derivative flow to compete for the channel effectively, the key idea is that the service tags for these flows in the participating nodes' flow tables should have correct relationship, i.e., if $T(f_1) \leq T(f_2)$ in one node's flow table, then it should be the same in the other node's flow table, so that nodes can make correct decisions in the backoff algorithm and the switch between sender-initiated and receiver-initiated handshake. It does not matter even if there are some discrepancies about the service tags of these flows maintained individually by each node.

In dealing with two-way flows, it is important to differentiate between original and derivative flows: The original flow is the one from the node that initiates the connection to the other node that acknowledges the connection. Then the required special processing can be

```
1: if (My flow is original) {
2:     if (Receive a data packet $T_d$ from the derivative flow)
3:             $f_{nd} = f_o + T_d$;
4: } else if (My flow is derivative) {
5:     if (Receive a data packet from the original flow)
6:             $f_{nd} = f_o$; $f_{no} = f_o + f_d$;
7: }
```

Figure 7.3: Special tag processing for two-way flows

summarized in two rules.

Rule 1: When a node that initiates the original flow receives a packet from the corresponding derivative flow, it sets the service tag for the derivative flow (maintained in its flow table) to be the service tag of the original flow plus the size of the acknowledged data packet measured in bytes. It does not change the *OtherFlow* flag because in fact the derivative flow is not an independent flow.

Rule 2: When a node that is the sender of a derivative flow receives a packet from the corresponding original flow, it updates the service tags for both flows in its table as follows. Let f_o denote the received service tag of the original flow and f_d the current service tag of the derivative flow in its table. Then the new service tags for the original flow (f_{no}) and the derivative flow (f_{nd}) are: $f_{nd} = f_o$ and $f_{no} = f_o + f_d$. In this case, the node does not change *MyFlow* flag because the node itself is in effect not making any real progress.

Figure 7.3 shows the algorithm when the above two rules are applied.

How to apply these rules are better illustrated by the example shown in Table 7.2. In this example, the packet from the original flow (0→1) has a size of 100 bytes, and the packet from the derivative flow (1→0) has a size of 4 bytes. *My* and *Other* are the short names for

Table 7.2: An example of two-way flow processing

Time	Event	Node 0				Node 1			
		$0{\rightarrow}1$	$1{\rightarrow}0$	My	Other	$0{\rightarrow}1$	$1{\rightarrow}0$	My	Other
t_0	initialization	0	0	-	-	0	0	-	-
t_1	0 sends data, 1 acks	100	0	1	-	0	0	-	-
t_2	1 sends data, 0 acks	100	104	1	-	0	4	1	-
t_3	0 sends data	100	104	1	-	100	4	1	-
t_4	1 acks	200	104	1	-	104	100	1	-

the *MyFlow* and *OtherFlow* flags. Rule 1 is used at time t_2 and rule 2 is used at time t_4. It is clear that these two rules make sure that the service tags of these two flows have the correct relationship in either node's table even if they are not up-to-date.

7.4 Simulation Results

In our simulations, we focus on how two competing flows share the available channel resource in a few simple network configurations. These configurations are shown in Figure 5.3. Despite the simpleness of these configurations, it is interesting to note that the fair schemes [14, 42, 84] proposed in the literature and the hybrid scheme proposed in Chapter 6 have not addressed all the fairness problems in these network configurations when flows are either UDP- or TCP-based.

We use GloMoSim 2.0 [73] as the network simulator and our implementation of the new scheme (TAFA) is based on the IEEE 802.11 MAC protocol. Below are some details of the implementation. For RTS/CTS, we add three fields: service tag (4 bytes), position flag (2 byte) of current flow and receiver-initiated (RI) flag (2 byte). Though 1 byte should be enough for any of these flags, we choose larger size to allow easy extensions in the future if any. For

Table 7.3: Common protocol configuration parameters

Data payload	DIFS	SIFS	
1460-byte	50μsec	10μsec	
contention window	slot time	sync. time	prop. delay
31–1023	20μsec	192μsec	1μsec

Table 7.4: IEEE 802.11 and TAFA specific configuration parameters

	RTS	CTS	data header	ACK
802.11	20-byte	14-byte	28-byte	14-byte
TAFA	28-byte	22-byte	48-byte	34-byte

data and ACK, in addition to the above three fields, they also include an advertisement about a flow from its flow table which includes three fields: source address (4 bytes), destination address (4 bytes) and service tag (4 bytes). In our implementation, a node indicates explicitly its originality in the RTS/data packets it sends out if applicable. All these constitute the fixed packet overhead in using the new scheme.

For IEEE 802.11, direct sequence spread spectrum (DSSS) parameters are used throughout the simulations. Configuration parameters that are common to all protocols are shown in Table 7.3 and protocol-specific configuration parameters are shown in Table 7.4. The raw channel bit rate is 2Mbps.

We investigate the performance of the IEEE 802.11 MAC protocol, the hybrid channel access scheme (for simplicity, it is simply called *Hybrid* thereafter) and the TAFA scheme under both UDP- and TCP-based traffic. In the first set of the simulation experiments, there are two competing UDP-based flows as described in Section 5.4. We have run each configuration five times with different seed numbers and with a duration of 30 seconds. If the standard

Table 7.5: Throughput comparison for the IEEE 802.11, the hybrid scheme and TAFA – two CBR flows (throughput measured in kbps)

Config #	Scheme	Flow #	Throughput	Flow #	Throughput	Aggregate
4-1	802.11	$0 \rightarrow 1$	83.4	$2 \rightarrow 3$	1500	1580
	Hybrid	$0 \rightarrow 1$	369	$2 \rightarrow 3$	1230	1600
	TAFA	$0 \rightarrow 1$	771	$2 \rightarrow 3$	778	1550
4-2	802.11	$1 \rightarrow 0$	820	$2 \rightarrow 3$	814	1630
	Hybrid	$1 \rightarrow 0$	-	$2 \rightarrow 3$	-	-
	TAFA	$0 \rightarrow 1$	769	$3 \rightarrow 2$	769	1540
4-3	802.11	$0 \rightarrow 1$	688	$3 \rightarrow 2$	709	1400
	Hybrid	$0 \rightarrow 1$	665	$3 \rightarrow 2$	643	1310
	TAFA	$0 \rightarrow 1$	683	$3 \rightarrow 2$	656	1340
4-7	802.11	$0 \rightarrow 1$	783	$2 \rightarrow 3$	824	1610
	Hybrid	$0 \rightarrow 1$	-	$2 \rightarrow 3$	-	-
	TAFA	$0 \rightarrow 1$	764	$2 \rightarrow 3$	764	1530
4-8	802.11	$0 \rightarrow 1$	1550	$3 \rightarrow 2$	28	1580
	Hybrid	$0 \rightarrow 1$	1280	$3 \rightarrow 2$	319	1600
	TAFA	$0 \rightarrow 1$	773	$3 \rightarrow 2$	805	1580
4-9	802.11	$0 \rightarrow 1$	734	$2 \rightarrow 3$	809	1540
	Hybrid	$0 \rightarrow 1$	815	$2 \rightarrow 3$	742	1560
	TAFA	$0 \rightarrow 1$	681	$2 \rightarrow 3$	676	1360

deviation of throughput is within 10% of the mean throughput, we show mean values only. Otherwise, we show both mean and standard deviation of the throughput. Table 7.5 shows the configurations when the IEEE 802.11 MAC protocol has fairness problem or there is some difference among these schemes. The "-" sign in the rows for the hybrid scheme indicates that receiver-initiated handshake is not triggered in these cases and it performs the same as IEEE 802.11 does. It can be seen that in some configurations such as 4-3 and 4-9 when the existing IEEE 802.11 MAC protocol works well, both the hybrid scheme and TAFA are unnecessary. Still, throughput degradation in TAFA is negligible. On the other hand, in configurations such as 4-1 and 4-8 where serious fairness problems occur in IEEE 802.11, TAFA shows superior performance to the other two schemes.

Table 7.6: Throughput comparison for the IEEE 802.11, the hybrid scheme and TAFA – two FTP flows (throughput measured in kbps)

Config #	Scheme	Flow #	Throughput	Flow #	Throughput	Aggregate
4-1	802.11	$0 \to 1$	0	$2 \to 3$	926	929
	Hybrid	$0 \to 1$	-	$2 \to 3$	-	-
	TAFA	$0 \to 1$	249	$2 \to 3$	423	672
4-2	802.11	$1 \to 0$	488±103	$2 \to 3$	453±102	942
	Hybrid	$0 \to 1$	439±99	$3 \to 2$	502±98	940
	TAFA	$0 \to 1$	383	$3 \to 2$	390	773
4-3	802.11	$0 \to 1$	530±432	$3 \to 2$	392±438	922
	Hybrid	$0 \to 1$	397±71	$3 \to 2$	455±78	852
	TAFA	$0 \to 1$	272	$3 \to 2$	363	635
4-7	802.11	$0 \to 1$	928	$2 \to 3$	0	930
	Hybrid	$0 \to 1$	-	$2 \to 3$	-	-
	TAFA	$0 \to 1$	443	$2 \to 3$	332	775
4-8	802.11	$0 \to 1$	929	$3 \to 2$	0	930
	Hybrid	$0 \to 1$	-	$3 \to 2$	-	-
	TAFA	$0 \to 1$	409	$3 \to 2$	209	617
4-10	802.11	$0 \to 1$	376	$3 \to 2$	526	902
	Hybrid	$0 \to 1$	-	$3 \to 2$	-	-
	TAFA	$0 \to 1$	335	$3 \to 2$	438	773

In the second set of simulation experiments, there are two competing TCP-based flows as described in Section 5.4. It should be noted that the acknowledgment packet from TCP is still regarded as a normal data packet from the view of traditional MAC layer, which does not provide special processing for two-way flows. However, in TAFA, the data flow and the acknowledgment flow are regarded as the original flow and derivative flow, respectively, and special processing is invoked as discussed in Section 7.3.4. Simulation results are shown in Table 7.6 for only the configurations when the IEEE 802.11 MAC protocol has fairness problems.

It is clear from Table 7.6 that the fairness problem is much more severe for two competing TCP-based flows than for the case of UDP-based flows if no special processing is

in place. For example, in some cases, such as configurations 4-1, 4-7 and 4-8, one FTP flow is denied access to the shared channel for most of the time. The hybrid scheme, due to the lack of flow contention information, cannot trigger the desired receiver-initiated collision avoidance handshake, hence it is of no avail. On the other hand, TAFA achieves much better fairness though at a cost of degraded throughput. This is a much desired tradeoff because it avoids the starvation of some flows and hence channel bandwidth is more evenly distributed among participating nodes. For configuration 4-3, please note the high variation of the throughput for these two flows in the case of IEEE 802.11 which shows that one flow monopolizes the channel for a long time and then gives it away to the other flow. Both the hybrid scheme and TAFA help to solve the problem. For other configurations, TAFA suffers some degradation in throughput. However, the overall performance of TAFA shows a much better tradeoff between throughput and fairness among the three schemes we investigate. We expect that even better algorithms than TAFA can be designed in the future following our fairness framework.

7.5   Conclusion

In this chapter, we have proposed a framework to address the fairness problem in ad hoc networks systematically. The framework includes four key components: Exchange of flow contention information, adaptive backoff algorithm, hybrid collision avoidance handshake, and special processing for two-way flows. We propose some specific algorithms to realize the framework and the resulting scheme, called topology aware fair access (TAFA), is evaluated through computer simulations against the IEEE 802.11 MAC protocol and the hybrid channel access scheme proposed earlier in Chapter 6. It is shown that TAFA can solve the fairness

problem in UDP-based applications with negligible degradation in throughput. TAFA is also quite promising for TCP-based applications, which have not been investigated at length in the past. Though TAFA suffers some throughput degradation, it solves the notorious problem of starvation of TCP flows, thus showing a much better overall tradeoff between throughput and fairness than the other schemes.

Given that the fairness framework is tailored to ad hoc networks and is general enough to accommodate new algorithms, it will be interesting to investigate new adaptive backoff algorithm and new criteria to switch between sender-initiated and receiver-initiated collision avoidance to achieve better throughput and fairness tradeoffs in future work.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

In this thesis, we present the first analytical model to derive the saturation throughput of four-way sender-initiated collision avoidance protocols in multi-hop ad hoc networks. We show that the sender-initiated collision-avoidance scheme achieves much higher throughput than the idealized carrier sense multiple access scheme with an ideal separate channel for acknowledgments. More importantly, we show that the collision avoidance scheme can accommodate much fewer competing nodes within a region in a network infested with hidden terminals than in a fully-connected network, if reasonable throughput is to be maintained. This shows that the scalability problem of contention-based collision-avoidance protocols looms much earlier than people might expect. Simulation experiments of the popular IEEE 802.11 MAC protocol validate the predictions made in the analysis.

Then we advance the aforementioned model to analyze MAC protocols that use di-

143

rectional antennas. These directional MAC protocols utilize different combinations of omni-directional and directional transmissions and receptions. Our advanced model consider both directional transmissions and receptions and the possible difference in gains between omni-directional and directional transmissions. Our analysis shows that, the scheme that uses a narrow antenna beamwidth in its transmissions and receptions achieves the best performance among all the schemes investigated. It also shows that when all-directional transmissions are augmented with directional receptions, one-hop throughput does not decrease due to the increased spatial reuse and reduced interference, even when the number of competing nodes within a region increases. This is very desirable because the scalability problem with the usual omni-directional collision avoidance schemes is mitigated at large by the use of powerful di-rectional antenna systems. It is also shown that, as expected, the performance of directional collision avoidance schemes degrades when directional transmissions have much higher gain than omni-directional transmissions because they cause interferences to more nodes that are originally outside the range of omni-directional transmissions but now are inside the range of directional transmissions. However, this degradation is relatively small when combined with directional receptions. Simulations of the IEEE 802.11 protocol and its directional variants validate the results predicted in the analysis; and show that side lobes affect little on through-put if the gain of the main transmission lobe is reasonably higher than that of side lobes and the carrier sensing threshold is raised to make nodes less sensitive to channel activities.

We also conduct a simulation study when there is a mix of unicast and broadcast traffic, as prior work only investigated the performance of broadcast traffic or unicast traffic using directional MAC schemes in isolation. Simulation results show that the presence of

144

broadcast traffic does not degrade the performance of the all-directional collision avoidance scheme significantly, even for relatively large percentages of broadcast traffic. The work indicates that an all-directional scheme is very attractive and practical for ad hoc networks. It attains much better throughput and delay than the other schemes, and the neighbor protocol that it needs to obtain location information of neighboring nodes can be implemented using very simple methods, without degrading its performance significantly. Our work also shows that the performance of broadcast traffic can be enhanced in a indirectly way because unicast traffic can be delivered with less delay in an all-directional scheme.

In addition to enhancing throughput in MAC protocols, alleviation of fairness problems is also very important. We investigate the fairness problem in detail and show the insufficiencies of prior backoff-based schemes and flow contention graph. We then propose a novel hybrid channel access scheme that combines both sender-initiated and receiver-initiated collision avoidance handshake. The new scheme is compatible with the popular IEEE 802.11 MAC protocol and involves only some additional queue management and book-keeping work. Simulation experiments show that the new scheme can alleviate the fairness problems existent for both UDP and TCP based applications with almost no degradation in throughput. The work also indicates the necessity of exchanging information among nodes so that nodes can compete for the channel more intelligently. Hence, we propose a framework to address the fairness problem more systematically. The framework identifies four key components for fair medium access. Based on the framework, we propose a topology aware fair access (TAFA) scheme to realize the framework. The novelties of the scheme include: adaptive backoff algorithm that use flow information, hybrid collision avoidance handshake that also depends

on topology information and special processing for two-way traffic. Simulations results show that TAFA can solve the fairness problem in UDP-based applications with negligible degradation in throughput. It can also solve the notorious problem of starvation of flows in TCP-based applications, despite some moderate degradation in throughput. Hence, TAFA shows a much better overall tradeoff between throughput and fairness than the other schemes investigated.

8.2   Future Work

As identified throughout the thesis, the following topics are interesting and worth further investigation.

First, we have shown that the all-directional MAC scheme can achieve high throughput and low access delay and the use of omni-directional transmissions to broadcast neighbor information is enough to obtain necessary information for directing antennas without degrading much throughput. However, it will still be interesting to evaluate neighbor protocols that can track other nodes' locations (or directions) and see how well the all-directional MAC scheme works in an environment where nodes can be moving in and outside one another's transmission and reception range.

Second, the enhancements to broadcasting proposed in the literature have not been incorporated in our investigation of directional MAC schemes, it will be interesting to investigate how the use of directional antennas as another dimension in the solution space can help to enhance broadcasting in multi-hop ad hoc networks.

Third, the fairness framework may still be enhanced in several ways. For example, in all the schemes discussed so far, the MAC layer serves requests in a first-in-first-out (FIFO)

manner, whether the request is an RTS to be sent for a data packet from upper routing layer, or an CTS to be replied to a neighbor in a sender-initiated or receiver-initiated mode. This may cause head-of-line (HOL) problems when the intended receiver faces severe contention and the HOL request cannot get through even though the requests behind the HOL packet may get through because their intended receivers are other nodes that face less contention. Hence, it may be advantageous for the MAC layer to maintain separate queues and achieve some prioritized or differentiated access. Other ways of enhancements include design of new adaptive backoff algorithm and new criteria to switch between sender-initiated and receiver-initiated collision avoidance to achieve better throughput and fairness tradeoffs.

Fourth, the work on fairness framework and mechanisms, though more inclined towards fair medium access among neighboring nodes, has already shown the benefits of more interactions between MAC layer and transport layer. It will be interesting to work on the integration of channel access with fair scheduling which can guard against misbehaved nodes and is a key component to provide QoS assurances. It is argued that such integration is necessary to provide services beyond best-effort and can facilitate the deployment of high profile applications in the future.

In addition to the above interesting topics, it is also worthwhile to investigate medium access control for sensor networks that are energy constrained and require scalable and efficient protocols because of the vast quantities of nodes involved [90]. Design of energy-efficient and scalable MAC protocols is indeed very challenging.

Bibliography

[1] "IETF Mobile Ad-hoc Networks (MANET) Working Group." `http://www.ietf.org/html.charters/manet-charter.html`.

[2] "NRL Mobile Ad Hoc Networking Page." `http://protean.itd.nrl.navy.mil/manet/manet_home.html`.

[3] D. B. Johnson and D. A. Maltz, "Protocols for Adaptive Wireless and Mobile Networking," *IEEE Personal Communications*, vol. 3, pp. 34–42, Feb. 1996.

[4] C.-K. Toh, *Wireless ATM and Ad-Hoc Networks*. Kluwer Academic Publishers, 1997.

[5] J. J. Garcia-Luna-Aceves, C. L. Fullmer, E. Madruga, D. Beyer, and T. Frivold, "Wireless Internet Gateways (WINGs)," in *Proc. of IEEE MILCOM '97*, Nov. 1997.

[6] Z. J. Haas and S. Tabrizi, "On Some Challenges and Design Choices in Ad-hoc Communications," *Proc. of IEEE MILCOM '98*, vol. 1, pp. 187–192, 1998.

[7] M. Corson, J. Macker, and G. Cirincione, "Internet-based Mobile Ad Hoc Networking," *IEEE Internet Computing*, vol. 3, no. 4, pp. 63–70, 1999.

[8] J.-P. Hubaux, J.-Y. Le Boudec, S. Giordano, and M. Hamdi, "The Terminode Project: Towards Mobile Ad-Hoc WANs," in *IEEE International Workshop on Mobile Multimedia Communications (MoMuC '99)*, pp. 124–128, 1999.

[9] "Mobile Information and Communication Systems." `http://www.terminodes.org/`.

[10] "Rice University Monarch Project: Mobile Networking Architectures." `http://www.monarch.cs.rice.edu/`.

[11] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications," in *Proc. of the Fall Joint Computer Conference*, pp. 281–285, 1970.

[12] L. Kleinrock and F. A. Tobagi, "Packet Switching in Radio Channels: Part I - Carrier Sense Multiple-access Modes and Their Throughput-delay Characteristics," *IEEE Trans. on Communications*, vol. 23, no. 12, pp. 1400–1416, 1975.

[13] P. Karn, "MACA - a New Channel Access Method for Packet Radio," in *ARRL/CRRL Amateur Radio 9th Computer Networking Conference*, pp. 134–140, ARRL, 1990.

[14] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A Media Access Protocol for Wireless LANs," in *Proc. of ACM SIGCOMM '94*, 1994.

[15] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Floor Acquisition Multiple Access (FAMA) for Packet-Radio Networks," in *Proc. of ACM SIGCOMM '95*, 1995.

[16] IEEE Computer Society LAN MAN Standards Committee, ed., *IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Std 802.11-1997, The Institute of Electrical and Electronics Engineers, New York, 1997.

[17] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Solutions to Hidden Terminal Problems in Wireless Networks," in *Proc. of ACM SIGCOMM '97*, 1997.

[18] F. Talucci and M. Gerla, "MACA-BI (MACA by Invitation): A Receiver Oriented Access Protocol for Wireless Multihop Networks," in *Proc. of PIMRC '97*, 1997.

[19] F. Talucci and M. Gerla, "MACA-BI (MACA by Invitation): A Wireless MAC Protocol for High Speed Ad Hoc Networking," in *Proc. of IEEE ICUPC '97*, pp. 913–917, 1997.

[20] J. J. Garcia-Luna-Aceves and C. L. Fullmer, "Performance of Floor Acquisition Multiple Access in Ad-Hoc Networks," in *Proc. of 3rd IEEE Symposium on Computers and Communications*, 1998.

[21] C. L. Fullmer, *Collision Avoidance Techniques for Packet-Radio Networks*. PhD thesis, University of California, Santa Cruz, CA 95064, U.S.A., June 1998.

[22] J. J. Garcia-Luna-Aceves and C. L. Fullmer, "Floor Acquisition Multiple Access (FAMA) in Single-channel Wireless Networks," *ACM/Baltzer Mobile Networks and Applications*, vol. 4, no. 3, pp. 157–174, 1999.

[23] J. J. Garcia-Luna-Aceves and A. Tzamaloukas, "Reversing The Collision-Avoidance Handshake in Wireless Networks," in *Proc. of ACM/IEEE Mobicom 1999*, (Seattle, WA, USA), Aug. 1999.

[24] J. J. Garcia-Luna-Aceves and A. Tzamaloukas, "Receiver-initiated Collision Avoidance in Wireless Networks," *ACM Wireless Networks*, vol. 8, pp. 249–263, 2002.

[25] L. Bao and J. J. Garcia-Luna-Aceves, "A New Approach to Channel Access Scheduling for Ad Hoc Networks," in *Proc. ACM Seventh Annual International Conference on Mobile Computing and networking*, (Rome, Italy), July 2001.

[26] C. Zhu and M. S. Corson, "A Five-Phase Reservation Protocol (FPRP) for Mobile Ad Hoc Networks," in *Proc. of IEEE INFOCOM '98*, vol. 1, (San Francisco, CA, U.S.A.), pp. 322–31, 1998.

[27] F. A. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part II - the Hidden Terminal Problem in Carrier Sense Multiple-access Modes and the Busy-tone Solution," *IEEE Trans. on Communications*, vol. 23, no. 12, pp. 1417–1433, 1975.

[28] H. S. Chhaya and S. Gupta, "Throughput and Fairness Properties of Asynchronous Data Transfer Methods in the IEEE 802.11 MAC Protocol," in *Proc. of PIMRC '95*, 1995.

[29] G. Bianchi, L. Fratta, and M. Oliveri, "Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs," in *Proc. of PIMRC '96*, pp. 392–396, 1996.

[30] I. Widjaja, B. P. Crow, L. G. Kim, and P. T. Sakai, "Investigation of the IEEE 802.11 Medium Access Control (MAC) Sublayer Functions," *Proc. of IEEE INFOCOM '97*, vol. 1, pp. 126–133, 1997.

[31] J. Weinmiller, M. Schläger, A. Festag, and A. Wolisz, "Performance Study of Access Control in Wireless LANs – IEEE 802.11 DFWMAC and ETSI RES 10 Hiperlan," *ACM/Baltzer Mobile Networks and Applications*, vol. 2, no. 1, pp. 55–67, 1997.

[32] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 535–547, Mar. 2000.

[33] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 1774–1786, Sept. 2000.

[34] F. Cali, M. Conti, and E. Gregori, "Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit," *IEEE/ACM Transactions on Networking*, vol. 8, pp. 785–799, Dec. 2000.

[35] Y.-B. Ko, V. Shankarkumar, and N. H. Vaidya, "Medium Access Control Protocols Using Directional Antennas in Ad Hoc Networks," in *IEEE INFOCOM 2000*, Mar. 2000.

[36] A. Nasipuri, S. Ye, J. You, and R. E. Hiromoto, "A MAC Protocol for Mobile Ad Hoc Networks Using Directional Antennas," in *Proc. of the IEEE Wireless Comm. and Networking Conf. (WCNC) 2000*, (Chicago, IL, USA), Sept. 2000.

[37] R. Ramanathan, "On the Performance of Ad Hoc Networks with Beamforming Antennas," in *ACM MobiHoc '01*, (Long Beach, CA, U.S.A.), Oct. 2001.

[38] M. Takai, J. Martin, A. Ren, and R. Bagrodia, "Directional Virtual Carrier Sensing for Directional Antennas in Mobile Ad Hoc Networks," in *ACM MobiHoc '02*, (Lausanne, Switzerland), June 2002.

[39] R. R. Choudhury, X. Yang, R. Ramanathan, and N. Vaidya, "Medium Access Control in Ad Hoc Networks Using Directional Antennas," in *ACM MobiCom '02*, (Atlanta, GA, USA), Sept. 2002.

[40] Z. Huang and C.-C. Shen, "A Comparison Study of Omnidirectional and Directional MAC Protocols for Ad Hoc Networks," in *IEEE Globecom 2002*, (Taipei, Taiwan), Nov. 2002.

[41] A. Nasipuri, K. Li, and U. R. Sappidi, "Power Consumption and Throughput in Mobile Ad Hoc Networks using Directional Antennas," in *Proc. of the IEEE Intl. Conf. on Computer Comm. and Networks (ICCCN)*, (Miami, Florida, USA), Oct. 2002.

[42] T. Ozugur, M. Naghshineh, P. Kermani, C. M. Olsen, B. Rezvani, and J. A. Copeland, "Balanced Media Access Methods for Wireless Networks," in *Proc. of ACM/IEEE MOBICOM '98*, pp. 21–32, Oct. 1998.

[43] N. H. Vaidya and P. Bahl, "Fair Scheduling in Broadcast Environments," Tech. Rep. MSR-TR-99-61, Microsoft Research, Dec. 1999.

[44] T. Nandagopal, T. Kim, X. Gao, and V. Bharghavan, "Achieving MAC Layer Fairness in Wireless Packet Networks," in *ACM Mobicom 2000*, (Boston, MA, USA), Aug. 2000.

[45] N. H. Vaidya, P. Bahl, and S. Gupta, "Distributed Fair Scheduling in a Wireless LAN," in *ACM Mobicom 2000*, (Boston, MA, USA), Aug. 2000.

[46] H. Luo, S. Lu, and V. Bharghavan, "A New Model for Packet Scheduling in Multihop Wireless Networks," in *ACM Mobicom 2000*, (Boston, MA, USA), Aug. 2000.

[47] H. Luo and S. Lu, "A Topology-Independent Fair Queueing Model in Ad Hoc Wireless Networks," in *IEEE ICNP 2000*, (Osaka, Japan), Nov. 2000.

[48] H. Luo, P. Medvedev, J. Cheng, and S. Lu, "A Self-Coordinating Approach to Distributed Fair Queueing in Ad Hoc Wireless Networks," in *IEEE INFOCOM 2001*, Apr. 2001.

[49] Y. Wang and B. Bensaou, "Achieving Fairness in IEEE 802.11 DFWMAC with Variable Packet Length," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM '01)*, (San Antonio, TX, U.S.A.), Nov. 2001.

[50] X. Huang and B. Bensaou, "On Max-min Fairness and Scheduling in Wireless Ad-Hoc Networks: Analytical Framework and Implementation," in *ACM MobiHoc '01*, (Long Beach, CA, U.S.A.), Oct. 2001.

151

[51] Y. Wang and J. J. Garcia-Luna-Aceves, "Collision Avoidance in Multi-Hop Ad Hoc Networks," in *Proc. of IEEE/ACM Intl. Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems (MASCOTS '02)*, (Forth Worth, TX, U.S.A.), Oct. 2002.

[52] Y. Wang and J. J. Garcia-Luna-Aceves, "Performance of Collision Avoidance Protocols in Single-Channel Ad Hoc Networks," in *Proc. of IEEE Intl. Conf. on Network Protocols (ICNP '02)*, (Paris, France), Nov. 2002.

[53] Y. Wang and J. J. Garcia-Luna-Aceves, "Modeling of Collision Avoidance Protocols in Single-Channel Multihop Wireless Networks," *ACM Wireless Networks Journal, Special Issue on Modeling and Analysis of Mobile Networks*, vol. 10, no. 5, 2004.

[54] Y. Wang and J. J. Garcia-Luna-Aceves, "Spatial Reuse and Collision Avoidance in Ad Hoc Networks with Directional Antennas," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM '02)*, (Taipei, Taiwan), Nov. 2002.

[55] Y. Wang and J. J. Garcia-Luna-Aceves, "Collision Avoidance in Single-Channel Ad Hoc Networks Using Directional Antennas," in *Proc. of IEEE Intl. Conf. on Distributed Computing Systems (ICDCS '03)*, (Providence, Rhode Island, U.S.A.), May 2003.

[56] Y. Wang and J. J. Garcia-Luna-Aceves, "Broadcast Traffic in Ad Hoc Networks with Directional Antennas," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM '03)*, (San Francisco, CA, U.S.A.), Dec. 2003.

[57] Y. Wang and J. J. Garcia-Luna-Aceves, "A New Hybrid Channel Access Scheme for Ad Hoc Networks," in *Proc. of the 1st Annual Mediterranean Ad Hoc Networking Workshop (Med-hoc-Net)*, (Sardegna, Italy), Sept. 2002.

[58] Y. Wang and J. J. Garcia-Luna-Aceves, "Throughput and Fairness in a Hybrid Channel Access Scheme for Ad Hoc Networks," in *Proc. of the IEEE Wireless Comm. and Networking Conf. (WCNC '03)*, (New Orleans, Louisiana, U.S.A.), Mar. 2003.

[59] Y. Wang and J. J. Garcia-Luna-Aceves, "Channel Sharing among Competing Flows in Ad Hoc Networks," in *Proc. of IEEE Symposium on Computers and Communications (ISCC '03)*, (Kemer - Antalya, Turkey), June 30 – July 3 2003.

[60] Y. Wang and J. J. Garcia-Luna-Aceves, "A New Hybrid Channel Access Scheme for Ad Hoc Networks," *ACM Wireless Networks Journal, Special Issue on Ad Hoc Networking*, vol. 10, no. 4, 2004.

[61] Y. Wang and J. J. Garcia-Luna-Aceves, "Modeling of Directional Collision Avoidance in Ad Hoc Networks," *Elsevier Performance Evaluation Journal*, 2004. To appear.

[62] Y. Wang and J. J. Garcia-Luna-Aceves, "Performance of Directional Collision Avoid-

ance in Ad Hoc Networks," in *Proc. of the Third IFIP-TC6 Networking Conference (Networking 2004)*, (Athens, Greece), May 2004.

[63] J. J. Garcia-Luna-Aceves and Y. Wang, "Throughput and Fairness of Collision Avoidance Protocols in Ad Hoc Networks," in *Ad Hoc Networks: Technologies and Protocols* (S. V. Krishnamurthy and P. Mohapatra, eds.), Kluwer Academic Publishers, 2004. To appear.

[64] H. S. Chhaya, "Performance Evaluation of the IEEE 802.11 MAC protocol for Wireless LANs," Master's thesis, Illinois Institute of Technology, Chicago, IL, U.S.A., May 1996.

[65] S. Chhaya and S. Gupta, "Performance of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol," *IEEE Pers. Comm.*, vol. 3, Oct. 1996.

[66] B. P. Crow, I. Widjaja, L. G. Kim, and P. T. Sakai, "Performance of IEEE 802.11 Wireless Local Area Networks," *Proc. of SPIE Conference on Digital Backbone and Fast Packet Technologies*, Nov. 1996.

[67] B. P. Crow, I. Widjaja, L. G. Kim, and P. T. Sakai, "IEEE 802.11 Wireless Local Area Networks," *IEEE Communications Magazine*, vol. 35, pp. 116–126, Sept. 1997.

[68] G. Bianchi, "IEEE 802.11–Saturation Throughput Analysis," *IEEE Communications Letters*, vol. 2, pp. 318–320, Dec. 1998.

[69] Y. C. Tay and K. C. Chua, "A Capacity Analysis for the IEEE 802.11 MAC Protocol," *ACM/Baltzer Wireless Networks*, vol. 7, no. 2, pp. 159–171, 2001.

[70] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. on Info. Theory*, vol. 46, pp. 388–404, Mar. 2000.

[71] H. Takagi and L. Kleinrock, "Optimal Transmission Range for Randomly Distributed Packet Radio Terminals," *IEEE Transactions on Communications*, vol. 32, no. 3, pp. 246–57, 1984.

[72] L. Wu and P. Varshney, "Performance Analysis of CSMA and BTMA Protocols in Multihop Networks (I). Single Channel Case," *Information Sciences, Elsevier Sciences Inc.*, vol. 120, pp. 159–77, 1999.

[73] X. Zeng, R. Bagrodia, and M. Gerla, "GloMoSim: a Library for Parallel Simulation of Large-scale Wireless Networks," in *Proc. of the 12th Workshop on Parallel and Distributed Simulations*, May 1998.

[74] S. Basagni, D. Bruschi, and I. Chlamtac, "A Mobility-Transparent Deterministic Broadcast Mechanism for Ad Hoc Networks," *IEEE/ACM Transactions on Networking*, vol. 7, pp. 799–807, Dec. 1999.

[75] E. Pagani and G. P. Rossi, "Providing Reliable and Fault Tolerant Broadcast Delivery in Mobile Ad Hoc Networks," *ACM/Baltzer Mobile Networks and Applications*, vol. 4, pp. 175–192, 1999.

[76] I. Chlamtac, A. Myers, V. Syrotiuk, and G. Zaruba, "An Adaptive Medium Access Control (MAC) Protocol for Reliable Broadcast in Wireless Networks," in *IEEE International Conference on Communications*, (New Orleans, LA, U.S.A.), June 2000.

[77] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network," in *Proc. of ACM/IEEE Mobicom 1999*, (Seattle, WA, U.S.), Aug. 1999.

[78] H. Lim and C. Kim, "Flooding in Wireless Ad Hoc Networks," *Computer Communications*, vol. 24, pp. 353–363, Feb. 2001.

[79] B. Williams and T. Camp, "Comparison of Broadcasting Techniques for Mobile Ad Hoc Networks," in *ACM MobiHoc 2002*, (Lausanne, Switzerland), June 2002.

[80] K. Tang and M. Gerla, "MAC Layer Broadcast Support in 802.11 Wireless Networks," in *Proceedings of IEEE MILCOM 2000*, (Los Angeles, CA, USA), Oct. 2000.

[81] K. Tang and M. Gerla, "MAC Reliable Broadcast in Ad Hoc Networks," in *Proceedings of IEEE MILCOM 2001*, (McLean, VA), Oct. 2001.

[82] K. K. Ramakrishnan and H. Yang, "The Ethernet Capture Effect: Analysis and Solution," in *Proc. 19th Conference on Local Computer Networks (LCN '94)*, (Minneapolis, MN, USA), Oct. 1994.

[83] K. Christensen, "Performance Evaluation of the Binary Logarithmic Arbitration Method (BLAM)," in *Proc. 21st Conference on Local Computer Networks (LCN '96)*, (Minneapolis, MN, USA), Oct. 1996.

[84] B. Bensaou, Y. Wang, and C. C. Ko, "Fair Medium Access in 802.11 Based Wireless Ad-Hoc Networks," in *IEEE/ACM Intl. Workshop on Mobile Ad Hoc Networking and Computing (MobiHoc '00)*, (Boston, MA, U.S.A.), Aug. 2000.

[85] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: the Single Node Case," *IEEE/ACM Trans. on Networking*, vol. 1, pp. 344–357, 1993.

[86] S. J. Golestani, "A Self-Clocked Fair Queueing Scheme for High Speed Applications," in *Proc. of INFOCOM'94*, 1994.

[87] J. C. R. Bennett and H. Zhang, "WF2Q: Worse-case Fair Weighted Fair Queueing," in *Proc. of IEEE INFOCOM '96*, 1996.

[88] P. Goyal, H. M. Vin, and H. Chen, "Start-time Fair Queueing: A Scheduling Algorithm for Integrated Service Access," in *Proc. of ACM SIGCOMM '96*, Aug. 1996.

[89] X. Yang and N. Vaidya, "Priority Scheduling in Wireless Ad Hoc Networks," in *ACM MobiHoc '02*, June 2002.

[90] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magazine*, pp. 102–114, Aug. 2002.